DTIC

Data Link Technology For A Portable Unmanned Aerial Vehicle

Victor Kowalenko, Jane Phipps and Keith Cameron

DEPARTMENT OF DEFENCE

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION

Data Link Technology For A Portable Unmanned Aerial Vehicle

Victor Kowalenko, Jane Phipps and Keith Cameron

Air Operations Division
Aeronautical and Maritime Research Laboratory

DSTO-RR-0087

ABSTRACT

This report examines data link requirements for a portable unmanned aerial vehicle. Crucial to the operation of such a data link is the development of suitable computer algorithms that are capable of significantly compressing and reconstructing image data in a timely manner for viewing at a remote station. As a consequence of the near real-time requirement, we investigate recent advances in lossy data compression techniques concentrating on transform coding techniques involving the discrete cosine transform, fractals and wavelets. At present the discrete cosine transform is available on a microprocessor chip and can offer acceptable reconstructed images close to real-time with compression ratios of up to 35:1, but other techniques promise even higher compression ratios and possibly a near real-time capability in the not too distant future.

DEPARTMENT OF DEFENCE

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION

Published by

DSTO Aeronautical and Maritime Research Laboratory
PO Box 4331
Melbourne Victoria 3001

Telephone: (03) 9626 8111
Fax: (03) 9626 8999

AR No. AR-009-761
November 1996

APPROVED FOR PUBLIC RELEASE

Data Link Technology For A Portable Unmanned Aerial Vehicle

Executive Summary

In this report we review the current state of the art regarding three of the most prominent image compression techniques, namely discrete cosine, fractal and wavelet transform coding techniques. Our interest in these techniques arises out of a desire to achieve as much compression as possible in the real time transmission of image data from the sensor payload onboard a portable Unmanned Aerial Vehicle (p-UAV) to a ground control station similar in form to a laptop computer. In particular, we find that compression ratios of greater than 30:1 are required in order to receive VGA images of 640X400 resolution and 24 bit colour at a TV frame rate of 25 Hz. This would be suitable for Line Of Sight (LOS) naval surveillance operations. Even greater compression ratios are required to receive images of 320X200 resolution and 8 bit grey scale at a frame rate of 1 Hz which would allow for transmission along a High Frequency (HF) data link in land based reconnaissance operations. An HF data link removes the need for LOS transmission, although it may be beyond the present capabilities of all three compression techniques to produce images of acceptable quality. For LOS transmission in the Very and Ultra High Frequency bands image compression would still be required, particularly if transmission were to occur at the TV frame rate. Even for lower frame rates image compression is desirable either to reduce the demand on the limited power source of the p-UAV, extend the vehicle's range of operation or make transmission more jam resistant.

Of the three lossy techniques, the most popular is the Discrete Cosine Transform (DCT) technique, which is based on discrete Fourier transform theory. Although still undergoing further development as described in this report, the DCT technique has already met the standard for image compression put forward by the Joint Photographic Experts Group (JPEG). The DCT is limited in the amount of compression that can be achieved without serious degradation of the data resulting in block artefacts appearing on reconstructed images. Thus we consider the non-conventional techniques of fractal and wavelet transform coding, which promise even higher compression ratios than the DCT.

Fractal transform coding relies on the fact that many real world objects possess local self-similarity and can be described in terms of fractal transformations. These can be transmitted along a communications channel using less bandwidth than the pixel data of the original digital image. Fractal images not only provide a resolution independent image of the original, but can also yield very high compression ratios. However, at present it is questionable whether fractal coding is feasible for real-time applications, this report covers recent research directed at this question. In particular, we discuss recent attempts aimed at reducing the time expended in searching the domain blocks for each range block of the original image during encoding. We also describe the

Accurate Fractal Rendering Algorithm which enables the fast decoding of video streams. These developments offer real hope that a fractal encoding/decoding system will be available for near real-time applications in the not too distant future.

Wavelets can be viewed as bumps that can be squeezed or expanded by dilation and shifted by translation. An arbitrary function can be decomposed into a series of wavelets forming a complete orthonormal set, the underlying principle behind wavelet transform coding. Wavelet coding has attracted much interest over the past few years, mainly because it can bring about a reduction in the block artefacts associated with the DCT. Thus, it promises better quality reconstructed images at higher compression ratios than the DCT. At present it is unable to match the real-time performance of the DCT and may never reach those of fractal transform coding. With further advances in microprocessor technology and in optimising the software approaches described in this report, there is more than a possibility that this technique can be applied to near real-time applications soon.

The value to Defence of this work is a greater understanding of the current state of lossy image compression techniques for possible implementation in communication systems where large amounts of data are required to be transmitted over narrow bandwidths.

Authors

Victor Kowalenko
Air Operations Division

Victor Kowalenko graduated with a B.Sc. (Hons) in physics from the University of Melbourne in 1978. Awarded a Commonwealth Post-Graduate Research Award he undertook a Ph.D. in theoretical quantum plasma physics at the same university, completing it in 1982. Thereafter, he joined the former Materials Research Laboratories to carry out research into the theoretical plasma physics of the railgun project. In 1987 he moved to the Aeronautical Research Laboratory, now part of AMRL, to undertake research into aircraft systems and remotely piloted vehicles. Between 1991 and 1994, he was on leave taking up one of the inaugural Australian Research Fellowships at the University of Melbourne where he undertook research into the role of quantum plasmas in astrophysics and in general relativistic magnetohydrodynamics. On his return to Air Operations Division at AMRL, he was involved primarily in employing operational research models to simulate military scenarios. Returning to the University of Melbourne in 1995, Dr Kowalenko has been carrying out research into magnetic field dissipation of accreting plasma in the vicinity of a black hole, and into the properties of magnetised quantum plasmas. He has also been conducting research in various fields of applied mathematics, including asymptotics, mathematical methods and classical analysis.

Jane Phipps
Air Operations Division

Jane Phipps obtained a B.Sc (Hons) in Pure and Discrete Mathematics from the University of Tasmania in 1991. She spent several years tutoring and lecturing a wide variety of mathematical subjects. During this time she also spent two years on a PhD researching Cauchy Singular Integral Equations at the University of Tasmania. She left this institute to join AMRL in Melbourne in 1995 where her work has been in Operations and Performance Analysis.

Keith Cameron
Air Operations Division

Keith Cameron obtained a B.Sc. (hon) in aeronautical engineering from the City University London while working for the UK Defence Department. He then studied at the Cranfield Institute of Technology and was awarded a M.Sc. in aerodynamics in 1975. Returning to Australia in 1975 he joined the Department of Transports Flying Unit as a flight operations engineer. He moved to the then Aeronautical Research Laboratory (ARL) in 1976 to partake in remotely piloted aircraft development. During his time at ARL he has worked on systems aspects of a number of projects including Ikara and Nulka. This work has include aerodynamic testing, development of data acquisition systems, autopilot development, trials and studies at both systems and operational levels. He is currently working in the area of instrumentation and data acquisition systems for aircraft trials and wind tunnels.

Contents

1. Introduction

In a previous report [1] Cameron and Kowalenko, hereafter referred to as CK, discussed the feasibility of a portable Unmanned Aerial Vehicle (p-UAV) for deployment in various close range reconnaissance and surveillance missions currently being conducted by the Australian Defence Forces (ADF). It was expected that such a system would provide a capability estimated to be 50 to 60% of the performance of much larger and significantly more expensive systems. Also discussed in the report was the need to employ data compression techniques when considering the transmission of realtime image data from the vehicle to the Ground Control Station (GCS). They pointed out that in order to transmit realtime TV pictures a large bandwidth was required, which meant that transmission could only occur in the Very High Frequency (VHF) and Ultra High Frequency (UHF) bands. Hence, the range of the p-UAV was restricted to Line Of Sight (LOS) operations.

In CK it was stated that a combination of reducing the frame rate and compressing the transmitted data would bring about a decrease in the bandwidth, thereby reducing the power requirements of the system substantially. Thus, it was proposed that in the short term the data link for the p-UAV could employ the conventional transform coding techniques to compress data in the VHF/UHF bands, which in turn would provide the p-UAV with an operational range of about 30 km using a directional antenna. A longer term goal might be to employ more novel data compression techniques offering even higher compression ratios such as fractal and wavelet transform coding combined with the reduced frame rate. This could either:

(a) extend the range of the p-UAV;

(b) remove LOS limitations by operating in the lower frequency HF band;

(c) reduce directional antenna requirements;

(d) improve jam resistance.

In this report we aim to investigate the current state of the art with regard to lossy data compression techniques being employed in the transmission of image data over high frequency channels. Although many transform coding techniques exist [2], we shall be concerned primarily with the Discrete Cosine Transform (DCT) technique, which has become the standard for the Joint Photographic Experts Group, more commonly known as JPEG, and the non-conventional fractal and wavelet transform techniques. Although the latter two promise higher compression ratios without as much visible degradation than the former technique in specific applications, they are still evolving and as a consequence, they have not as yet replaced the DCT as the principal image compression technique. In particular, as we shall see, it is only due to advances in the last few years that fractal transform coding has been able to offer the possibility of a near realtime capability, which is so crucial in receiving image data from a UAV.

Realtime fractal transform coding is currently receiving much attention and we aim to discuss these developments in the present report.

The contents of this report are arranged as follows. Section 2 contains a summary of the basic principles of data transmission, which are necessary for understanding why transmitted images need to be compressed. The next section discusses the basic theory of electromagnetic propagation required selecting an appropriate carrier frequency for the transmission of data between the p-UAV and the GCS. In Section 4 we present basic information theory that not only clarifies the need for employing data compression techniques, but can also be used to evaluate them. Section 5 contains a description of the DCT both from a theoretical and practical point of view. In Secs. 6 and 7 we describe the current state of the art regarding the non-conventional techniques of fractal and wavelet transform coding. In discussing the three data compression techniques we relegate the mathematics to three separate appendices. Section 8 concludes with an evaluation of the techniques in regard to the p-UAV.

2. Modulation Techniques

Data links are not only required in all UAV systems for the transmission of realtime data but also for the navigation and control of the vehicle. Data link requirements which may include the need for data compression are determined primarily by the rate of data transmission between the vehicle and the GCS. Specifically, an uplink is required for manoeuvring the vehicle via the GCS whereas two downlinks are required respectively for monitoring the vehicle's position and for the transmission of image data collected by the sensor payload on the vehicle. The second downlink is referred to as a wideband downlink since the transmission of image data requires a much larger bandwidth than that required for the narrow links connected with navigation and control of the vehicle. Typical transmission rates for uplink control signals are less than 10 kHz while those for the transmission of sensor payload data, may require a transmission rate greater than 10 MHz depending on mission requirements for the UAV.

In a telecommunication system the first requirement is that the original information energy is converted or modulated into electronic signals [3]. These signals may then require amplification to increase the power levels before they are transmitted to a receiver at the destination. On reception the signals may be amplified again before being converted or demodulated into recognisable replicas of the original information. Thus, a complete system [4] consists of:

(a) a transmitter, (which includes the source of the original information),

(b) the transmission medium, and

(c) a receiver.

In order to transmit data efficiently, data transmission rates or more specifically, the sinusoidal electromagnetic signals, known as carrier signals, are modulated by superimposing on them information signals. Common forms of modulation are;

(a) Frequency Modulation (FM),

(b) Amplitude Modulation (AM),

(c) Phase Shift Key (PSK),

(d) Frequency Shift Key (FSK),

(e) pulse modulation which includes amongst others Pulse Amplitude Modulation (PAM), and

(f) Pulse Code Modulation (PCM).

In addition, variants of these modulation methods exist, for example, Differential PCM (DPCM), Quadrature or Quaternary PSK (QPSK) and Differentially-encoded QPSK (DQPSK or 4-phase DPSK). The last technique is efficient whilst allowing reliable reception with a simple demodulator [5].

FM and AM are essentially continuous wave modulation techniques, which involve analogue information signals with the former conserving transmitter power better than the latter [6]. However, here we concentrate on the transmission of digital signals rather than analogue signals since there are definite advantages in adopting a digital mode of transmission. In particular, digital signals are robust and can be easily processed [7] and because of their regenerative property, they can be transmitted over long distances through multiple switching centres and relay links with little noise interference or impairment. In addition, they are easily multiplexed, switched or recorded. Thus, a digital data link will be less affected by noise, which, in turn, means that less power is required to transmit signals with the same bandwidth. Furthermore, since CK proposed that the sensor payload consist of a CCD camera combined with a second or third generation image intensifier in Ref. [1], compatibility with such a sensor payload is easier to achieve by employing a digital transmission system rather than an analogue transmission system.

The remaining modulation techniques given in the above list operate on digital or binary data, although pulse modulation (PM) systems require some form of c-w modulation in transmitting data [4]. In a PCM system signal information is transmitted in digital form by sampling an analogue signal at regular intervals to produce a pulse amplitude modulated signal. Therefore, a PCM system can utilise solid state digital components [8]. Consequently, PCM tends to be favoured in applications where mass, cost and power consumption need to be minimised.

Before proceeding further with our evaluation of modulation techniques, we need to consider the display of images at the GCS. Although capable of offering higher resolutions at cheaper prices and of producing images with significantly less glare, Cathode Ray Tube (CRT) screens are more bulky, operate at higher voltages, consume more power and are less robust than Liquid Crystal Displays (LCDs). For these reasons LCDs are used in laptop computers. They also come in different dimensions with a variety of resolutions, the most common being 320X200, 640X400 and 640X480. In order to display images with adequate detail, more levels of grey (or colour) are required for the low resolution screens than for the higher resolution screens. Thus the number of colours required to obtain adequate imagery for a resolution of 320X200 would be at least 256 whereas for a resolution of 640X480 perhaps only two colours (black and white) are required since this resolution is superior to that for newspapers. A more detailed discussion of the relative merits of both CRTs and LCDs is given in Ref. [9].

FSK and PSK are commonly used in the transmission of data along computer and printer communication lines or cables. Since our aim is to produce digitised images on a laptop LCD (one of the requirements for the p-UAV presented in Ref. [1]), variants of FSK and PSK become the preferred modulation techniques. For covert transmission the preferred modulation technique is a PSK variant because it provides a constant amplitude, thereby permitting the power density to be spread evenly over the entire extended electromagnetic spectrum [10]. One disadvantage of coherent PSK systems, however, is that the receiver requires a good phase reference, which is difficult to achieve in practice. Thus the signal is degraded by extracting a phase reference from the transmitted signal.

The frequency spectrum of a modulated signal is usually symmetric about the carrier frequency except for SSB transmissions. The combined parts of the spectrum above and below the carrier frequency form the bandwidth of the signal. For the case where sensor data are transmitted in TV format the bandwidth depends on the number of picture elements per frame and the rate of transmission [8]. One important aspect in transmitting large amounts of data over a single communication link is the provision of pre-allocated channels that are capable of passing high enough frequencies, i.e. channels of large enough bandwidth. Commercial TV channels require about 6 MHz, but the bandwidth required for the transmission of a signal is dependent on the type of modulation employed by the telecommunication system.

As the modulated signals traverse the transmission medium or communication channel, they always become distorted due to additive noise (contamination by unwanted signals) and interference, which, in turn, place limitations on the transmission of data. At the receiver the best possible replica of the original information signal or message is obtained by demodulation, i.e. the removal of the carrier signal from the modulated signal and then filtering to remove noise.

3. Propagation

Because received messages are not perfect replicas of the original message, there is always a certain amount of uncertainty involved in decoding the received data. This, however, can be reduced by increasing the power of the original signal. Thus the ratio of the average signal power to the noise power (SNR) is an indication of the uncertainty or error in the received data. An increase in bandwidth allows a signal to be converted into a form that makes it more immune to noise. To transmit a given amount of data, either the signal power can be increased and the bandwidth reduced or the signal power can be reduced and the bandwidth increased.

The energy radiated by a transmitter can reach the receiving station or GCS by using one or more of the following modes [3,11]:

(a) surface or ground waves, which propagate by following the earth's curvature to distances of 15 to 110 km depending on frequency. An appreciable amount of the energy of electromagnetic waves is dissipated in this mode. These waves have frequency bands of below 300 Hz, 300 Hz-3 kHz, 3-30 kHz, 30-300 kHz and 300 kHz-3 MHz and 3-30 MHz, which correspond to Extremely Low Frequency (ELF), Infra Low (ILF), Very Low Frequency (VLF), Low Frequency (LF), Medium Frequency (MF) and High Frequency (HF) waves.

(b) sky waves, which are used for HF radio communications systems including long distance radio telephony and sound broadcasting. HF sky waves can be used on different portions of the frequency range 3-30 MHz at different times of the day. Sky waves are directed into the ionosphere and under certain conditions can be reflected to the required destination. The maximum and minimum useable frequencies for transmission change during the day. Hence, an operating frequency must be chosen from this range, which places a further restriction on the bandwidth for the transmission of signals in the HF band. For example, at 1400 hr the difference between the two frequencies is about 16 MHz whereas at about midnight the difference is about 2 MHz. In addition, interference from other users and atmospheric noise contribute even further to channel limitations. For more detailed information about the propagation characteristics of HF waves, see Ref. [12].

(c) space or LOS waves, which are utilised for both sound and television broadcasting and operate in the VHF (Very High Frequency), UHF (Ultra High Frequency) and SHF (Super High Frequency) bands of 30-300 MHz, 300 MHz-3 GHz and 3-30 Ghz respectively. These waves are dependent upon the distance to the horizon.

(d) via satellite systems such as the GPS system described in the previous section. Satellite systems can be employed to carry multi-channel telephony systems, TV signals and data in the UHF and SHF bands.

(e) scatter systems, which operate in the UHF and SHF bands. These are employed in multi-channel telephony links.

For a discussion of electromagnetic propagation problems in tactical environments such as battlefields and aircraft systems the reader is referred to Ref. [13].

In data transmission, a binary pulse is a pulse with one of only two possible amplitudes or states whilst a binary message or signal is a sequence of binary pulses occurring at regularly spaced intervals of say 1/R sec or at a rate of R/sec [4]. A bit is defined as the maximum amount of information that can be transmitted in a single binary pulse [14]. The uplink radio control for a UAV requires a mean data rate of up to 1 kbit/s and that a similar transmission rate for the telemetry data on the downlink is required for the necessary flight and management functions [6].

All telecommunication systems are capable of transmitting a maximum number of bits per second without loss in a channel. This is known as the channel capacity and is the central concept of data communication [15]. In addition to bandwidth, the channel capacity is limited by the Signal to Noise Ratio (SNR) for the system [14]. If a channel were free from noise, then the channel capacity would be infinite but since there is always some noise in the process of transmission, the channel capacity is always finite. For SNRs greater than 1 the channel capacity exceeds the bandwidth assuming that the noise is Gaussian and white. To cite interesting examples of channel capacity, the maximum downlink rate for the Space Shuttle is 48 Mbit/s while that for the proposed Space Station Freedom is 75 Mbit/s [2].

Current UAV transmission frequencies are restricted to an upper limit of 15 GHz whereas the lowest frequency when using an omnidirectional antenna is 2 MHz [10]. In a hostile ECM environment the effective jamming zone for a UAV can be reduced significantly, if the UAV is equipped with a directional antenna, but this may require a higher transmission frequency. There are several factors in selecting an appropriate carrier frequency for a data link, which, in turn, influence the cost, mass and power consumption of a UAV [8]. The minimum detectable signal power at a receiver is inversely proportional to the square of the frequency and hence, a lower frequency is more desirable. However, for a fixed antenna size, increasing the frequency results in increased gain, although the beamwidth is narrowed. Lower frequencies are also preferred to minimise the cost and mass of link components and atmospheric loss increases with increasing frequency. For naval applications vertically polarised radiation is preferred since multipath effects are greater for horizontally polarised radiation. Different carrier frequencies separated by several signal bandwidths should be used for both the uplink and downlink.

Although an HF data link provides much greater flexibility in that there is no need to maintain a line of sight with the vehicle, there are two major problems in creating such a data link for the PUMA. The first problem is the limited allocated bandwidths or channels for the transmission of data in the HF spectrum. These HF channels, which are often referred to as voice channels, are 3 kHz wide. By using higher powered

transmitters to provide SNRs greater than 3 dB, the transmission rate can be extended to 4.8 kbits/s. With the limited power available on the p-UAV, however, it is most likely that the channel capacity will be limited to 2.4 kbits/s (a quarter of the standard transmission rate for computer lines). For an HF data link to be viable, at a reduced frame rate, a transmission rate in the vicinity of 10 kbits/s is needed.

Single Side Band (SSB) transmission systems can use up to 4 voice channels (12 kHz bandwidth), two above the carrier frequency and two below [12]. Therefore, it is feasible to consider wideband HF data links, although in practice, problems may be encountered in obtaining approval to use more than one voice channel. However, as the p-UAV is likely to be deployed in missions conducted in remote areas over distances up to 30 km from the GCS [1], this approval may not be so difficult to obtain.

The second problem in using HF waves is concerned with their propagation characteristics. Sky waves are particularly useful as a means of transmitting signals over distances greater than 150 km, and may achieve ranges of over 3,200 km, although predicting the reflection off the ionosphere and obtaining consistent long range communication can be difficult. For short distances up to 80 km the groundwave mode is the appropriate form of HF communication, although certain conditions such as manpack radio operations in dense wet terrain can limit the usefulness of this mode to only a few km [16]. Furthermore, the gap beginning where the groundwave becomes too weak for communication and ending where the sky wave returns to earth has been considered as a region where HF communication is ineffective and is referred to as the skip zone, which in dense mountainous terrain can range from 4 to 150 km and hence, includes the operating radius of the p-UAV.

The problem described in the previous paragraph can be overcome by directing an HF skywave signal within a narrow band of frequencies at the zenith and then receiving the reflected wave back on Earth from one of several of the ionised layers in the atmosphere with a minimal path loss. This propagation mode is referred to as the Near Vertical Incidence Skywave (NVIS) mode and has been used by the US army since World War 2. The NVIS mode can be used to eliminate skip zones by adjusting antenna heights and transmitter frequencies [17]. The best frequency of operation for this mode lies in the 2 to 10 MHz frequency band [18]. The mode is also dependent on the directivity and polarisation of the antenna and any ground wave present can cause interference effects. Nevertheless, an HF-SSB radio with modern features operating in the NVIS mode can be used successfully to provide satisfactory communications for low flying tactical aircraft over a 50 km (or greater) range in virtually any type of terrain condition [18]. During daytime operations a lower power output can be used although at night and during dawn a high power output must be used to overcome the presence of noise and interference. With a 400 W transmitter the probabilities of communications success is 1.00, 0.92 and 0.87 for operation during the day, at dawn and at night respectively whereas with a 40 W transmitter these probabilities drop to 0.87, 0.59 and 0.73 respectively [18].

The NVIS mode must be considered for the following situations [16] when:

(a) the area of operations is not conducive to ground wave communications such as mountainous terrain,

(b) tactical deployments that place stations in anticipated skip zones when using whip antennas, frequency selection methods and operating procedures,

(c) operating in dense wet vegetation or other areas of high signal attenuation,

(d) prominent terrain features are not under friendly control,

(e) operating from defiladed positions,

(f) operating against enemy groundwave jammers and direction finders, and

(g) flying close to the ground.

The above indicates that it is feasible to design an HF data link utilising either ground waves or NVISs, although the bandwidth will be very small in comparison with that from a data link using carrier frequencies in the VHF, UHF or SHF bands. However, the UHF and SHF links are limited to maintaining a LOS with the p-UAV. Another problem with data links using higher frequency carrier waves is that they require costly and heavy components for operation and use significantly more power than their low frequency counterparts in the HF, MF and lower bands. Furthermore, if VHF or higher frequency carrier waves are used for data communication, then the GCS may require a directional antenna to track the vehicle as in the case of Pointer UAV [19].

The selection of an appropriate data link for a UAV is dependent on the nature of the mission, the amount of power the vehicle can provide and the vehicle's size and design in accommodating the antenna. For example, an LOS data link would be able to provide continuous moving pictures during surveillance operations of river banks and nearby areas (currently being carried out by Regional Force Surveillance Units [20]) or in the Protection of Vital Assets (PVA) such as the surveillance of air field perimeters. However, in missions where the vehicle is flying in mountainous or dense terrain, it may be more appropriate to use HF data links even with their restricted bandwidths. In addition to the amount of power which can be supplied by the vehicle, the antenna dimensions for both the vehicle and the GCS must be considered. For example, both these factors limit TV transmission to about 80 km. To transmit beyond this distance, relay stations must be used, which simply receive the signal, amplify it and then retransmit it.

4. Data Compression

Data compression can be defined as the collection of techniques which reduce the amount of digital data carrying useful information. These techniques are essential for the efficient handling of digital information. In addition, they may be useful in limiting the effectiveness of jamming and may provide a means of transmitting several channels over the bandwidth of a communication link where normally only one uncompressed channel could be employed to transmit data. Image data differ from other forms of data in that they are noisier and hence cannot be preserved exactly, although they can be preserved sufficiently for the human visual system not to notice.

If digital TV images consisting of 512X512 picture elements or pixels requiring 6 bits/pixel are to be displayed on a screen at 25 frames per sec, then this would correspond to a transmission rate of at least 40 Mbit/s. Because the transmission rate is determined by bandwidth, it is also affected by the type of modulation technique. For example, for 3 bit codes 6 Mbit/s are sufficient to transmit a 1 MHz video signal on a DPCM system compared with 14 Mbit/s for a PCM system [21]. Furthermore, a pulse or signal that can assume n distinct states or levels carries information equal to the logarithm of the number of choices or simply $\log_2(n)$ bits, for example, an octal pulse (one with eight different voltage levels) can be represented by combinations of 3 bits for each level.

To convert a video signal into a digital one, it must be sampled and quantised. Sampling the amplitude of a modulating signal must be carried out at regular time intervals. According to the Nyquist criterion, if a waveform has a bandwidth of f Hz, then it is possible to convey all the information in that waveform by 2f or more equally spaced samples per second of the amplitude of the waveform [22]. In practice a sampling rate of at least 4 times the video bandwidth is required for each signal [23]. Quantisation, on the other hand, is the assignment of approximate discrete intensity values for the amplitudes of the sampled points. The best contrast performance for the lowest number of coded bits per sample is achieved by using 3 bits corresponding to eight intensity levels. Thus for a video bandwidth of 2.3 MHz, which corresponds to the transmission of images with a resolution of 400X300 pixels at a frame rate of 20 Hz, a typical sampling rate would be 9.2 MHz and thus, a bit rate of at least 27.6 Mbps would be required assuming a three bit code (eight grey levels) for each sample [23].

Data messages are particularly susceptible to instantaneous loss of signal since fading or corruption by noise can result in the loss of a few bits that can destroy the information content. The most important characteristic of a data link is its minimum Bit Error Rate (BER) as opposed to the minimum received Signal to Noise Ratio (SNR) for an analogue channel. The difficult part in constructing a data link is minimising BERs by attempting to reduce noise power independently of signal power because there is a limit on the amount by which the latter can be increased. Frost et al [8] state that because the uplink of a UAV should respond correctly to the GCS commands, the BER should be as small as possible ($\approx 10^{-5}$) whereas BERs as high as 0.001 may be

acceptable in the display of image data due to the high level of redundancy. The amount of redundant data in a TV image has been estimated to be as high as 99.9% in special cases. The large amount of redundancy is attributed to the fact that the Human Visual System (HVS) responds most responsively to scene details of high contrast, i.e. the edges of objects in an image. BER values of 0.001 may be achievable using ground or sky waves but for the NVIS mode and the skip zone the BERs are likely to be much worse for considerable periods.

The ability to measure the maximum amount of information per second that a system can transmit cannot answer the question of which system or group of systems has sufficient capacity to transmit a specified class of information bearing messages. To answer this question the information content of a signal must be determined. For example, the appropriate system to transmit a speech in English is determined by the information content of the speech and the time available to complete transmission [4]. The information content in messages consisting of equally likely symbols is given by $I = M \log_2(N)$ where I is the information content, M is the number of symbols and N is the number of bits per symbol. However, in most cases certain letters and combinations of letters occur more often than others, for example, *e* occurs more frequently than *z* and *u* is more likely to appear after a *q*. Thus, the information content of a message not only relates to the number of possible signal combinations but also to their relative frequency of occurrence, which, in turn depends upon the source of the message. The measure of the amount of information contained in a set of data yields the entropy of the information source producing the data. Entropy is defined as the sum over all members of a symbol alphabet in which each probability of occurrence is multiplied by its logarithm to base 2.

For a digital video image, the symbols are the quantised intensity values at the pixels. For example, in 8-bit quantisation (typical for video quality animation) there are 256 symbols and the entropy is given in bits per pixel (bpp), which is known more commonly as the bit rate. In general, the intensity values at nearby pixels are highly correlated and as a consequence, digital video images contain much redundant data, thereby yielding a lower entropy.

The need for data compression in a p-UAV wideband uplink can now be demonstrated by considering the transmission of image data. For images of 320X200 pixels with 256 (8 bits) levels of grey on an LCD screen, a minimum of 0.51 Mbits/s are required for a new frame every second. Additional bits are required to indicate the start and end of each quantised level, the so-called start and stop bits. Therefore, 0.64 Mbits/s need to be transmitted to exhibit one frame per second. Furthermore, the actual rate may be even greater if an additional bit is required for error bit checking, the so-called error bit. However, the pre-allocated bandwidth for HF frequencies is about 10 kHz. If the assumption is made that the entire bandwidth can be used for the transmission of data, then to obtain 1 image per second on an LCD for an SNR of about 1, i.e. a channel capacity of 10 kbits/s, the data would need to be compressed by at least a factor of 64. The compression ratio is even greater for the transmission of monochrome (black and white) images with a 640X400 resolution.

For naval vessels conducting surveillance operations over an open sea, mission requirements may call for highly resolved images, for example, 600X400 pixels (VGA quality) with 24-bit colour at TV frame rates of 25 Hz. In this situation an LOS data link would be required. The given image resolution may be necessary since much of the information displayed on the images received at the GCS would consist of open sea and sky. On the other hand, Rejman [24] states that the effects of reducing frame rate are;

(a) manual navigation and feature-flying tasks take much longer to accomplish,

(b) task areas are covered less efficiently, and

(c) the task of target detection may be performed poorly.

Thus, the frame rate should be kept as high as possible even though in some cases reducing it may be necessary to accommodate the transmission of image data along a narrow pre-allocated bandwidth such as an HF data link.

To transmit VGA images, which only allow 16 colours (4-bits) per pixel, a transmission rate of 25 Mbits/s is required for a frame rate of 25 Hz. This transmission rate exceeds greatly the channel capacity of TV channels whose bandwidths are typically 5 MHz. For a system with an SNR of 10 dB, which is indicative of the minimum for a digital system, the channel capacity for a bandwidth of 5 MHz would be 17 Mbits/s. Thus, VGA images need to be compressed if the frame rate is kept at 25 Hz.

Even if the frame rate were reduced so that the amount of data transmitted to the GCS could be accommodated by a TV channel, there are other important reasons for compressing sensor payload data, provided that the resulting images do not exhibit serious degradation. If the system is to operate in an environment rendered hostile by the use of Electronic Counter Measures (ECM) such as jamming, then a reduction in signal bandwidth would reduce the vulnerability in the transmission of sensor data. In addition, the lower the signal bandwidth, the lower the power required to transmit data. Thus the transmission of less redundant information can be exploited to produce an effective improvement in SNR, i.e. reducing the bit rate is equivalent to increasing transmitter power. This is particularly important for UAVs with limited power sources such as the p-UAV investigated by CK [1].

There are numerous data compression techniques for compressing binary data [2], which can be categorised as follows;

(a) reversible or information-lossless image compression. Here the original digital representation of an image can be fully reconstructed at the receiver from the compressed data. Examples include run-length coding, contour coding, Huffman coding, arithmetic coding and conditional replenishment.

(b) predictive methods. These involve predicting the intensity value at a given pixel based on the values of previously processed pixels. Examples include the above-mentioned DPCM, Delta Modulation (DM), and Motion Compensation (MC).

(c) block methods. Here the image is subdivided into blocks, which are then processed in a variety of methods. For example, in Vector Quantisation (VQ) the blocks are compared to a codebook of vectors and the code with the *closest match* is transmitted while in block truncation coding the value at each pixel in a block is coded as a 0 or a 1 depending on whether it is above or below a chosen threshold. Digital data such as character strings cannot be transmitted using vector quantisation because small changes in the numerical value of a character lead to enormous changes in meaning.

(d) Human Visual System (HVS) compensation. These techniques attempt to compress video images by eliminating data not perceptible to the HVS, even if the data are important from an information theory point of view. Some techniques apply a model of the HVS directly to the image data whereas others have been developed to represent as many features of the HVS as possible. Examples include the method of synthetic highs, pyramid coding, regional growing and directional decomposition.

(e) transform coding. This information lossy technique uses a mathematical operator to produce an array of uncorrelated or nearly uncorrelated data from the highly correlated data representing a digital image. Examples include the Karhunen-Loeve Transform (KLT), the Discrete Cosine Transform (DCT), the Slant Transform and the Hadamard Transform;

(f) hybrid techniques. These consist a mixture of the techniques described above. An example is the DCT/VQ, which involves using VQ on the DCT coefficients.

Performance of data compression techniques is evaluated in terms of the Mean Square Error (MSE) or variants of it. The MSE is essentially an error measure consisting of the sum of the square of the intensity differences between the reconstructed and original images divided by the square of the number of intensity values. In later sections we shall use a variant, the Peak Signal-to-Noise Ratio (PSNR). This is given by,

$$\text{PSNR} = 20 \log_{10}(b/d_{rms}),$$

where, b is the largest possible value of the signal (typically 255) and d_{rms} is the root mean square error difference between two images f and g defined as

$$d_{rms} = \sqrt{\int_{I^2} \left(f(x,y) - g(x,y)\right)^2 \, dx\, dy} \qquad (1)$$

In the above, (x,y) refers to the position coordinates of both images.

Lossless techniques cannot offer the high compression ratios required for the p-UAV data link and of the remaining techniques the one that has become the most prominent is the DCT. This technique offers both high compression ratios and a low MSE [2], which the other transform techniques cannot offer. At the same time it meets the JPEG standard. However, there are two relatively new transform coding techniques which promise even higher compression ratios and good fidelity. These are fractal and wavelet transform coding techniques. In what follows we aim to discuss these techniques in detail and review the current state of the art in employing these techniques in image data compression.

5. The Discrete Cosine Transform

We begin our study of the image data compression techniques mentioned at the end of the previous section with the DCT, which was first employed by Ahmed et al [25]. The DCT is an orthogonal transformation in that mathematical operators are used to form a complete orthogonal set of unique basis vectors. The transform acts to 'pack' a large number of highly correlated image data samples into a smaller number of uncorrelated coefficients [2]. Of the three techniques, which we aim to review, the DCT algorithm is the only one that meets the JPEG standard for sequential lossy compression according to page 219 of Ref. 26 and as a consequence, it has become synonymous with JPEG compression.

A major advantage of the DCT is that its basis vectors are known. Hence these do not need to be calculated for every transform block, thereby reducing the encoding time. They also do not need to be transmitted together with the coefficients, thereby reducing the transmission time. Another advantage of the DCT is that there are already several fast algorithms for computing them [2].

DCT image compression involves dividing the original image into smaller NxN blocks and then transforming these blocks via the Forward Discrete Cosine Transform (FDCT) into equal-sized blocks of coefficients in the frequency domain. Because it employs the same basis vectors for each transform block, they only need to be evaluated for the first transform block with a lookup table being used for the other blocks. Data compression is achieved by assigning fewer bits to the coefficients in order to remove redundant information via a couple of methods. First, threshold sampling is used so that all coefficients above a certain magnitude are retained while those below the threshold value are set equal to zero. Second, the NxN array is compressed further by undergoing rounding-off or quantisation of the pixel intensity levels. The degree of quantisation is greater for the higher frequency coefficients since the human eye is more sensitive to rounding-off at the lower frequencies. The resulting data are encoded via a lossless technique such as arithmetic or Huffman coding to avoid the loss of time experienced in transmitting the many coefficients that become zero after quantisation. The data can then be transmitted over a communication link

with some error correcting code to enable decoding at the receiver by applying the Inverse DCT (IDCT) which gives a representation of the original output image [2].

When the DCT is implemented using the JPEG standard, the image is first partitioned into 8x8 blocks and the FDCT is applied to these [27]. In JPEG compression, thresholding and quantisation occur together in one matrix. The DC term, a DCT coefficient representing the mean pixel value for each block, is differenced from the DC term of the preceding block in a scanning order and the remaining coefficients are passed to an entropy encoder. An entropy coding scheme, typically Huffman coding, is then employed to assign codewords to coefficients in such a way that short codewords are assigned to the more frequent terms while longer ones are assigned to the rarer terms. The values are encoded in a zigzag manner as there is a high correlation between values along this zigzag scan. Decompression is accomplished by applying the inverse of each step in the opposite order [26,28]. For a discussion on the mathematical details concerning the implementation of the DCT the reader is referred to Appendix A.

5.1 Improvements in the DCT

JPEG compression can produce undesirable blocking artefacts for high compression rates. That is, if the amount of thresholding and quantisation is too severe, then sub-block boundaries may appear in the reconstructed image. At the same time setting too many high frequency coefficients to zero can lead to a loss in resolution. As a consequence, much activity is being directed at developing non-standard methods of employing the DCT and using JPEG compression as the bench-mark. Rather than rounding off to the nearest integer values after dividing by quantisation coefficients as in JPEG compression (see Appendix A), Eude et al [29] have recently proposed using a mixture of Gaussian distributions on DCT coefficients in their search for a better means of quantisation. They found that by approximating the high frequency DCT coefficients by a single Gaussian distribution and the low ones by a mixture of two or three Gaussian distributions, they were able with their new quantisation matrix to remove blocking effects present in JPEG compression.

Another area where DCT image compression is being improved is in the acceleration of the algorithmic process. For example, in Ref. [30] Hung and Meng describe two methods for accelerating the computation of the inverse DCT (see Appendix A) by exploiting the sparseness of the quantised transform coefficients. One method referred to as the Symmetric Mapped Inverse DCT (SMIDCT) can perform up to three times faster than the Forward Mapped Inverse DCT (FMIDCT) [31], the previous best optimisation of the inverse DCT for sparse matrices [32]. In addition, Jung and Mitra [33] have developed a method that not only reduces the blocking effect mentioned above, but also accelerates the computation of DCT coefficients. Basically their method involves decomposing the computation of an N-point DCT into a computation of an (N/2)-point DCT and an (N/2)-point Discrete Sine Transform (DST). Jung and Mitra refer to this method as SubBand DCT (SB-DCT) decomposition and have tested the

method on images with a 256x256 resolution. They have demonstrated that their SB-DCT method not only matches JPEG coding with respect to PSNR but also performs at least twice as fast. In addition, the compressed images exhibit much less blocking effects than the corresponding JPEG compressed images.

Khataie and Soleymani [27] have proposed two different two-stage image compression schemes aimed at achieving better quality images than JPEG compression for moderate to high PSNRs. In the first stage both schemes process the more important low frequency components of the image through transform coding while the high frequency components lost in the first stage are encoded in the second stage. The first scheme employs a DCT algorithm combined with a high rate Lattice-based Vector Quantiser (LVQ) algorithm in the first stage while the second scheme employs a standard JPEG encoder. For a description of LVQ schemes, most of which employ the algorithm designed by Linde, Buzo and Gray (known as the LBG algorithm) [34], the reader is referred to Ref. [2]. In the second stage each scheme processes the "error" of the residual image formed by subtracting the output of the first stage from the original image. Khataie and Soleymani use the same low-rate LVQ in both schemes. They find for PSNRs of greater than 38 dB the first scheme is much superior to JPEG compression. For example, they achieve 1.2 bpp for a PSNR of 40.56 as opposed to 1.8 bpp via JPEG compression. They also find that although 0.7 bpp can be achieved at a PSNR of 36.0 using JPEG compression compared with 1.05 bpp from the second scheme, the latter outperforms JPEG compression significantly for high quality compression. For example, 1.6 bpp can be achieved for a PSNR of 47.73 via the second scheme whereas JPEG offers only 3.9 bpp. Khataie and Soleymani conclude that both schemes offer a considerable improvement over standard JPEG compression for moderate to high quality compression. By using either scheme more than ninety five per cent of input information can be retrieved with a bit rate less than 2 bpp. However, these authors do not discuss the amount of time involved in processing images via both schemes. In addition, although these techniques are still under development, for them to be viable for the p-UAV much higher compression ratios are required.

In regard to reducing the processing time Walmsley et al [28] have proposed a pruning algorithm in which a smaller proportion of DCT values are calculated. That is, instead of calculating DCT values for an 8X8 image block, they only find it necessary to calculate the DCT values for a 4X4 subset whilst simultaneously maintaining an acceptable image quality. Their pruning algorithm requires a total of 82 multiplications and 227 additions for a block compared with 192 multiplications and 464 additions using the standard row-column approach. Another advantage is that parallelisation can be performed during stages of the algorithm because two or more processors can be invoked to calculate separate data partitions that arise from the decomposition of the DCT. When applied to the JPEG standard the pruning algorithm not only accelerates image compression as a result of calculating less DCT coefficients but also produces higher compression ratios with negligible degradation in image quality due to the fact that there is only one long length of zeros along the zig zag scan of the encoder. Walmsley et al find that for an acceptable loss in image quality, i.e. a pruning value of 4, their algorithm results in a speed up of over 50% on JPEG compression.

Another area of much interest is the introduction of new architecture designed to use less area of each processing chip. For example, Wang and Chen [35] have proposed using systolic arrays in DCT computation, which have received much attention ever since they were introduced by Kung and Leiserson [36] in the design of high speed signal processing systems. Systolic systems possess the desirable features of regularity, modularity and concurrency, thereby enabling parallel computing architectures to be created from them which are necessary in meeting the realtime requirements for the transmission of image data. On the other hand, Mariatos et al [37] have introduced a novel architecture employing the Coordinate Rotation Digital Computer (CORDIC) Circular Rotation Algorithm. The CORDIC algorithm is based on the decomposition of the DCT matrix into rotations. The new architecture requires less than 40% of the area of previous CORDIC architectures to perform DCT computations and when extensively pipelined (up to 80 pipeline stages can be set) it can process the fast signals of High Definition Television (HDTV). Mariatos et al have adopted 2-bit digit-serial arithmetic to bring about a reduction in hardware. The chip can perform at a throughput rate of 500 MHz or 250 Mpixels/s and needs about 2.6K gates.

Although the DCT has become the most popular image compression technique mainly because of its implementation in the JPEG standard, from the preceding material it can be seen that there is still room for improvement, especially in the transmission of HDTV signals or in the transmission of image data over relatively narrow bandwidths with a requirement for a near realtime capability such as our p-UAV application. Current activity is concerned mainly with reducing the processing time although improvements in the quantisation process may lead to marginally higher compression ratios. The techniques mentioned above should be monitored as they will undoubtedly lead to a new JPEG standard in the future. For our p-UAV application, however, we require much higher compression ratios than those offered at present by the DCT, particularly if we wish to consider HF propagation. Two techniques promising higher compression ratios than the DCT are fractal and wavelet transform coding and we shall investigate the current state of the art of these techniques in the following sections.

6. Fractal Transform Coding

A fractal is a fragmented geometrical shape that can be continually subdivided into parts, in which each part is a copy of the original shape only reduced in size. That is, fractals are generally self-similar. Many real world objects, which are not simple geometric shapes such as clouds, mountains and coastlines, can be described by fractals because real world images possess local self-similarity as described in Ch. 1 of Ref. 38. This means that only parts of images possess the same self-similar transformations and hence, an image consists of properly transformed parts of itself. These transformed parts seldom combine to form an exact copy of the original image.

As a consequence, an image encoded as a set of fractal transformations will not be an identical copy, but an approximation, i.e. lossy.

It should also be noted that while we are primarily concerned with realtime applications of fractal image compression in this report, the implementation of this technique in non-realtime environments has already met with remarkable success. In 1992 Microsoft released a compact disc known as *Microsoft Encarta*, which is a popular multimedia encyclopedia containing 7 hr of sound, 100 animations, 800 colour maps and more than 7000 pictures all encoded in less than 600 Mb of data. Microsoft has been able to achieve this astonishing feat using fractal image compression techniques. Thus, it is only a matter of time before consumers will use this technology to store their valuable pictures on compact disc rather than adopting the archaic procedure of storing them in photographic albums.

Deterministic fractals possess the intrinsic property of extremely high visual complexity while being very low in information content [39]. This is because they can be generated by simple recursive deterministic algorithms and it is this property that makes them a useful tool in image compression. Transformations used in the description and pointers to image regions are stored rather than the original pixel image data. Thus, fractal transform coding yields a set of relations based on the spatial and spectral geometry of the original image, which describe the original image in terms of itself. Fractal images not only provide a resolution independent image of the original, but can also yield very high compression ratios [40]. For a description of the mathematical concepts leading to the implementation of fractals in image compression the reader is referred to Appendix B. Here we shall be concerned more with the recent advances employing this transform coding technique.

We now describe how fractal image compression can be applied to an 256X256 image in which each pixel is any of 256 (8 bit) levels of grey. Let $R_1, R_2, \ldots R_{1024}$ be the 8X8 pixel *nonoverlapping* sub-squares of the image and let **D** be the collection of all 16X16 pixel (overlapping) subsquares of the image which yields 58,081 squares. For each range block R_i a search is conducted through all the domain blocks of **D** to find the block D_i which minimises the rms metric given by,

$$d_{rms}(f \bigcap (R_i \times I), w_i(f)) \quad i = 1, \ldots N . \tag{2}$$

That is, we find pieces D_i and maps w_i, so that when a w_i is applied to the part of the image over D_i, something very close to the part of the image over R_i is obtained. There are eight ways to map one square onto another, which means that $8 \bullet 58{,}081 = 464{,}648$ squares with each of the 1024 range squares. In addition, a square in **D** has four times as many pixels as each R_i, so that either subsampling, i.e. choosing one pixel from each 2X2 subsquare of D_i or averaging the 2X2 subsquares corresponding to each pixel of R_i must be adopted to minimise the above equation.

Minimising the rms metric requires not only finding a D_i of the image that looks most like the image above R_i but also finding good contrast and brightness settings s_i and o_i for the PIFS transformation w_i discussed in Appendix B. For each $D \in \mathbf{D}$, s_i and o_i are computed by using least squares regression, which also gives a resulting rms difference. The chosen D_i is the $D \in \mathbf{D}$ with the least rms difference.

The selection of D_i together with the corresponding s_i and o_i allows the transformation w_i to be put in matrix form. Once a collection $w_1, w_2, \ldots w_{1024}$ has been determined, the image can be decoded by estimating the attractor A as defined by Equation. (B3) in Appendix B. In general, not that many iterates are required to obtain a representation of the original image as Fisher shows on p. 15 of Ref. [38]. Here, representations are presented after the first, second and tenth iterates with the final one displaying all the essential features of the original 65,536 byte image. The transformations that reconstitute the image require only 3968 bytes since each transformation requires 8 bits in the x and y directions to determine the position of D_i, 7 bits for o_i, 5 bits for s_i and 3 bits to determine a rotation and flip operation for mapping D_i to R_i. The position of R_i is implicit in the ordering of the transformations. Hence, a compression ratio of 16.5:1 is obtained with an rms error of 10.4. Each pixel is on average only 6.2 grey levels from their correct value while with each iteration more detail is added.

According to Jacquin [39] the three main issues involved in the design of a fractal block coding system are;

(a) the partitioning of an image,

(b) the selection of a measure of the distortion between two images,

(c) the specification of both a finite class of contractive image transformations defined with a partition and of a scheme for the quantisation of their parameters.

In the remainder of this section we shall primarily be concerned with issues (a) and (c) while issue (b) is discussed in Appendix B.

6.1 Image Partitions

In the previous subsection we presented most of the ideas of a practical fractal image encoding scheme. An image is first partitioned by some collection ranges R_i and then for each R_i a domain block D_i from the collection of image pieces is sought that has a low rms error when mapped to R_i. Once R_i and D_i are known, s_j and o_j of a partitioned Iterated Function System (IFS) can be determined in addition to

a_j, b_j, c_j, d_j, e_j and f_j of the affine transformation. Eventually, a transformation $W = \bigcup w_j$ is obtained that encodes an approximation to the original image. So far, we have concentrated on fixed-size range blocks R_i, but there are regions in the original that are difficult to cover using this approach, for example, a person's eyes in a photograph. Furthermore, there are regions that can be covered with a larger R_i, thereby reducing the total number of maps w_i required. Optimal partitioning of an image is not only capable of improving the quality of the reconstructed image, but can also increase the compression ratio.

There are several methods of partitioning an image, some of which we describe here. First, Jacquin [39] presents a partitioning technique where an image is partitioned by using a block coding design based on the theory of iterated contractive image transformations. The original image μ_{orig} is partitioned into domain cells and into nonoverlapping square range cells of two different sizes forming a two-level square partition. A partition constructed in this way is image-dependent, although it does allow for the use of larger blocks to take care of smoothly varying image regions and smaller ones to capture detail in intricate regions such as rugged boundaries and fine textures. The domain blocks form the pool **D** consisting of all image blocks and these are then classified according to their features as;

(1) shade blocks, D_S,

(2) edge blocks, D_E, and

(3) midrange blocks, D_M.

Shade blocks do not possess significant gradients and are not used as domain blocks. Hence, they can be removed from the pool. On the other hand, edge blocks possess strong changes of intensity and are split further into simple and mixed edges. Midrange blocks possess moderate gradients but no definite edges.

Now consider an r×r digital image μ quantised to 256 grey levels. The original image μ is partitioned $\{R_i\}_{0 \le i < N}$ into range cells of two different sizes. The image transformation can be represented as;

$$\tau = \sum_{i=1}^{N-1} g_i \text{ with } g_i = T_i \circ S_i, \tag{3}$$

where S_i and T_i are the geometric and massic parts of g_i. First, the spatial construction S_i must be constructed by selecting an image domain block $\mu \urcorner_{D_i}$ of size D×D, which will be contracted to a block $S_i(\mu \urcorner_{D_i})$ of size B×B. The symbol $\urcorner_D$ represents the part of the image constrained to block D_i. The specification of the domain cell D_i is equivalent to the description of the spatial contraction S_i. The second part consists of finding the

block transformation T_i which minimises the distortion between $T_i \circ S_i(\mu\rceil_{Di})$ and $\mu\rceil_{Ri}$. The distortion measure used by Jacquin is the rms distortion between image blocks. A pool of massic transformations T can now be obtained. The encoding of the range blocks $\mu\rceil_R$ consists of utilising the self-transformability by finding the best pair $(D_i,T_i) \in D \times T$ such that the distortion $d(\mu\rceil_{Ri}, T_i \circ S_i(\mu\rceil_{Di}))$ is a minimum. By implementing this partitioning scheme Jacquin was able to achieve a bit rate of 0.06 bpp with a Peak SNR (PSNR) of 31.4 dB for an 8 bit 512X512 resolution of the image of *Lena*.

Jacquin's method, however, expends too much time because of the large amount of searching in the domain block pool. To overcome this problem, Bani-Eqbal [41] has devised a new technique for speeding up the search. He proposes an incremental method that employs Jacquin's method, but limits the domain block pixels by averaging to half their size. This is referred to as decimating [39]. He then flips them and compares them with the range blocks. The domain blocks are arranged into a tree, so that the tree can be navigated to select a small number of candidate blocks. By using this method he is able to achieve for a 256X256X8 bit version of *Lena* a speeding up of more than 50 times on the Jacquin's complete search method without any noticeable degradation in image quality. Specifically, he finds that it takes 8750 s to encode the image on a SUN Sparcstation 10 Model 30 using Jacquin's method whereas with his method it only takes 150 s with a marginal increase in rms distortion (6.7 for the complete search method as opposed to 8.7 with his method).

6.2 Quadtree Partitioning

Fisher et al [42] were first to introduce adaptive methods in the encoding process. They employed various approaches including quadtree, rectangular and triangular partitions of the range blocks to improve the fidelity of an image. They also pointed out that it is not necessary to impose strict contractivity conditions on each of the coded transformations since the eventual contractivity of their union is sufficient to ensure convergence of the iteration process during decoding.

In quadtree partitioning a square is divided into four equal sub-squares when it is not covered sufficiently by some domain. The process continues recursively beginning with the entire image and continuing until the squares are small enough to be covered within a specified rms tolerance. Small squares can be covered better than large ones because adjoining pixels in an image tend to be highly correlated. Thus, an image can be represented as a tree in which each node contains four subnodes, corresponding to the four quadrants of the square while the root of the tree is the initial image.

An algorithm based on the above method can be developed by assuming that the image contains 256X256 pixels. The collection of permissible domains D can be all the sub-squares of 8X8, 12X12, 16X16, 24X24, 32X32, 48X48 and 64X64. Next the image is partitioned recursively until the squares are 32X32. Then an attempt is made on each

square to be covered in a quadtree partition by a larger domain. The success of each attempt is determined by meeting a predetermined tolerance value e. When this condition is met, the square can be called R_i and the covering domain D_i. If the condition is not met, then the square is subdivided and the process repeated.

According to Fisher [38] the algorithm works well, but can perform even better if the domain pool includes diagonally oriented squares. As an example, Fisher states that for a 256X256 grey image of a collie the quadtree scheme yields a compression ratio of 28.95:1 with an rms error of 8.5. However, the domain-range comparison step of the encoding is computationally intensive and so a classification scheme is invariably used to minimise the number of comparisons.

Several classification schemes exist. One of these schemes is the block method used by Jacquin [39], as already described, while another is archetype classification [38]. Here an archetype A is determined by searching through the entire domain set to find that member of the set that covers the other members best. Here covering means that both the domain and corresponding transformation which result in an accurate mapping to each range are found. Archetype classification is similar to determining a Vector Quantisation (VQ) codebook, but a major difference is that the transformation, w, is included in the process of determining archetypes.

In Chapter 4 of Ref. [38] Boss and Jacobs introduce an archetype classification scheme which is subsequently employed to encode the standard *Lena* image. They describe how five different sets can be generated, each consisting of 72 archetypes. Three of the archetype sets are determined from three sets containing five qualitatively dissimilar 256X256X8 images while the remaining two are determined from sets of five qualitatively similar images. None of the sets of images contains the Lena image or any other test image. Boss and Jacobs show that when the number of classes from each archetype set is the same as conventional block classification schemes [43], the latter are able to encode images much faster than the archetype method, although image fidelity or rather, the PSNR, is significantly better using the former method. As a consequence, the number of archetype classifications can be lowered, which not only yields a better PSNR, but also reduces the encoding time to below that for the conventional scheme with its higher number of searched classes. By considering only six searched classes for the archetype method Boss and Jacobs show that the Lena image takes about 200s to encode on an Apollo 4500 workstation yielding a PSNR of 24.25 whereas with 24 searched classes in their conventional scheme encoding of the *Lena* image takes over 300 s yielding a PSNR of 24.1.

6.3 HV-Partitions

A deficiency in quadtree partitioning is that no attempt is made to select the domain pool D in a content-dependent manner. The selected collection must be very large to enable a good fit to the given range. A technique to overcome this deficiency whilst

simultaneously increasing the flexibility of the range partition is to employ HV-partitioning. In an HV-partition a rectangular image is recursively partitioned either horizontally or vertically to form two new rectangles until a covering tolerance is satisfied as in quadtree partitioning. This technique is much more flexible since the position of the partition is variable, thereby allowing the partitions to share some of the self-similar structure. For example, partitions can be arranged so that edges in the image will run diagonally through them. It is then possible to use the larger partitions to cover the smaller ones with the expectation of obtaining a good cover. For a more detailed description and variation of this technique the reader is referred to Chapter 6 of Ref. [38].

When encoding with HV-partitions, the same two basic steps required are as for the quadtree method. These are;

(1) recursive partitioning to establish nonoverlapping ranges, and

(2) domain searching to determine the domain that will map onto a particular range.

Each pixel in the original image is assigned to exactly one range through partitioning, but it can appear in multiple domains, which are typically two to three times greater than the ranges. As before, the affine transformation of the pixel values must be contractive.

In HV-partitioning, however, the average pixel value for each row and column of pixels, of the particular range undergoing partitioning, is calculated. These averages are used to compute successive differences between the averages. Then a linear biasing function is applied to each of these differences, which multiplies them by their distance from the nearest side of the rectangular range. That is, if the range contains pixel values $r_{i,j}$ for $0 \le i < N$ and $0 \le j < M$, then the horizontal sums,

$$\sum_i r_{i,j}$$

and vertical sums,

$$\sum_j r_{i,j}$$

are computed, subtracted and multiplied by the biases,

$$\min(j,M\text{-}j\text{-}1)/(M\text{-}1) \text{ and } \min(i,N\text{-}i\text{-}1)/(N\text{-}1),$$

respectively. The first partition is found by determining the maximum value of all the biased horizontal and vertical differences so that it is either located at horizontal position j or at vertical position i, depending on which yields the larger biased difference. This yields two rectangles which tend to partition the given range along the strong vertical or horizontal edges while avoiding narrow rectangular partitions.

The domain search is almost the same as the quadtree method. Once a rectangle is divided, a domain is sought for the largest currently uncovered range. Unlike the

quadtree method, it is not the rms difference but the square difference of the pixel values that is compared with a predetermined threshold to determine when partitioning takes place. If the square difference is smaller than the threshold, then the transformation is accepted, otherwise the range is partitioned into two new ranges.

The encoding time can also be accelerated by employing one of the following methods;

(1) quadrant classification,

(2) encoding by range size, or

(3) domain-range ratio restriction.

Decoding can be performed by a more efficient method than the standard method of iteration to a fixed point. This more efficient method involves pixel referencing and low-dimensional fixed point approximation. For more details of the above methods, the reader is referred to the article by Fisher and Menlove in Chapter 6 of Ref. [38].

The Fisher and Menlove technique can be applied at various optimisation levels ranging from 0 to 8 with each level resulting in a decrease in the number and type of comparisons to be performed. At high optimisation levels very rapid compression is achieved yielding high compression ratios, but not very good fidelity. Specifically, Fisher and Menlove obtained the following results;

Table 1 Results obtained by Fisher and Menlove

Level	Compression Ratio	PSNR (dB)	Encoding Time (secs)
5	14.7	34.62	979.9
2	39.4	30.99	1122.1
6	80.2	28.15	42.5

By inspecting the reproduced images they were able to rule out the final case as an acceptable level. In addition, they found that the relationship between the PSNR, compression ratio and encoding time is linear on a log scale.

Popescu and Yan [44] have also developed an adaptive block splitting scheme which is more flexible than a quadtree method. The image is split into blocks according to a tree structure with the base or roots consisting of the initial 24×24 block splits of the image. Branches are formed by two partitioning attempts; the first creates nine 8×8 blocks and the second four 12×12 blocks. On the first branch each 8X8 block is searched for a match in the pool of domain blocks. If none is found, then further branches are created by splitting the 8×8 block into four 4×4 blocks. On the second branch each 12X12 block is searched for a match in the pool of domain blocks and if none is found, then the next

level is investigated which consists of four 6×6 blocks, one 8×8 and five 4X4 blocks and nine 4×4 blocks. The path producing the shortest code is eventually selected. Popescu and Yan state that this splitting strategy produces optimal results compared with quadtree partitioning and have applied their method to a colour image of a fish achieving a high compression rate of 31.44 and a PSNR of 36.14, but the encoding time is not given.

6.4 The Bath Fractal Transform

Monro and Dudbridge [45] have developed the Bath Fractal Transform (BFT) method of encoding rectangular grey-scale image blocks which eliminates the need for searching. Zero searching fractal transforms are particularly important because both the coding and decoding speeds are fast. The image is tiled with reduced copies of itself using a least-squares approximation to derive an optimal mapping or set of affine transformations known as a Self Affine System (SAS). The approximation to a rectangularly tiled block is found by evaluating various low order moments over the block and solving a set of four linear equations for each tile. The method is easy to implement and is feasible for real-time applications. A brief description of the method appears below while more extensive details can be obtained from Refs. [45-47].

To encode an image, an Iterated Function System (IFS) of order N in R^2 is found. This is the SAS, which is defined as,

$$W = \{w_k : \mathrm{k} = 1,\ldots,\mathrm{N}\}, \tag{4}$$

where,

$$w_k = \begin{pmatrix} a_{11}^{(k)} & a_{12}^{(k)} \\ a_{21}^{(k)} & a_{22}^{(k)} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} b_1^{(k)} \\ b_2^{(k)} \end{pmatrix}.$$

The IFS is defined arbitrarily with a rectangular attractor and is called the domain part of the BFT. The attractor A could be, for example, any non-overlapping tiling of the image. For each k a fractal function $f(x,y)$ known as the function part of the transform, is defined on A which approximates the grey scale $g(x,y)$ of tile k. The fractal function is specified by a recursive set of mappings,

$$f(w_k(x,y)) = v_k(x,y,f(x,y)). \tag{5}$$

Contracting an image fragment $f(x,y)$ onto the image introduces self-similarity and hence, the process can be viewed as fractal. The mappings form a collection of functions which, when iterated or rendered, form an approximation to the image according to the Collage Theorem (see Appendix B). The BFT finds the least squares mapping and v_k can be any function that is contractive of f. When v_k is a polynomial in (x,y), the BFT involves evaluating low order moments over the image blocks and the solution of small sets of linear equations. Grey scale mappings can be represented as

$$v_k(x,y,f) = a + b_x x + b_y y + c_x x^2 + c_y y^2 + d_x x^3 + d_y y^3 + e f(x,y). \quad (6)$$

A zero order fractal is one where all its coefficients except for a and e are equal to zero. The first order terms in x and y are referred to as a bilinear fractal transform while the second order terms are referred to as biquadratic.

By minimising the Collage Theorem with respect to the coefficients, a fractal least squares approximation can be obtained to a given function $g(x,y)$. That is, by using the rms metric (Eq. (1)), one minimises for each k by taking partial derivatives of,

$$\int_{(x,y)\in A} \left[g(w_k(x,y)) - a - b_x x - b_y y - e\, g(x,y) \right]^2 dL, \quad (7)$$

and setting them equal to zero to obtain a solution for a, b_x, b_y and e.

Surprisingly, Monro and Woolley [47] have found that to obtain high fidelity or low compression images, it is better to employ higher order methods while low order fractal transform methods are better for low fidelity images. This means that the type of fractal transform method one employs is dependent upon the application. For example, if the p-UAV were to be employed in surveillance missions involving large expanses of water, then low order fractal transform methods may be suitable, whereas for reconnaissance missions in dense vegetation, the opposite would seem to apply.

Monro and Dudbridge [48] have also developed the Accurate Fractal Rendering Algorithm (AFRA) that enables fast decoding of video streams. The algorithm overcomes problems associated with traditional methods which attempt to construct an exact fractal when only a representative/finite set of pixels is required on a graphics screen of finite resolution. Determining this finite set is called rendering of the fractal. Monro and Dudbridge introduce the non-iterative Minimal Plotting Algorithm (MPA) to show how deterministic fractals can be rendered by generating a pixel set that approximates the minimum cover of an attractor. Compared with the Random Iteration Algorithm (RIA), the MPA is able to plot 101,258 points of the particular fractal known as the Sierpinski triangle using 303,774 transformations whereas only 84% of the MPA points are plotted with the RIA after 303,774 iterations. The AFRA is an adaptation of the MPA, which renders fractal functions as given by Equations (5) and (6). That is, it approximates grey scale images by a simple extension of the MPA. Monro and Dudbridge conclude that the MPA and AFRA display IFS fractals at any desired resolution in very few computations per pixel and that they help overcome a major barrier to the application of fractal technology by supporting real-time performance video compression/decompression. This is demonstrated in Ref. [48] where the BFT and AFRA are combined to produce a real-time fractal video compression scheme with bit rates as low as 40 kbits/s while still displaying images at 25 frames per second.

6.5 Parallel Processing

An alternative approach of avoiding the intense computations required in the fractal encoding of images is the massively parallel implementation scheme developed by Xue et al [49] for a multi-SIMD quad pyramid machine. Typically, fractal encoding complexity is $O(n^4)$ for an n×n image, which prohibits real-time application. The scheme in Ref. [49], however, reduces the complexity to $O(n^2)$, the same order as the decoding complexity. Using a 256X256 image, Xue et al achieved the following results on a pyramid machine with a base of mxm processors;

Table 2 Results achieved by Xue et. al.

m	B	τ_p (s)
256	8	32
	16	87
128	8	128
	16	350
64	8	513
	16	1399

In the table, B represents the number of bits in the data and τ_p is the parallel computing time. The scheme can be implemented on smaller machines, but as can be seen from the table, the computing times increase.

6.6 Fractals and the DCT

Sloan [50] has carried out two studies to compare fractal transform coding with JPEG implementations of DCT compression. The first study involved images with a resolution of 640X400 at 24 bits/pixel while the second involved images with a resolution of 1024X1024 at 16 bits/pixel. Image fidelity was measured by comparing the rms difference between the original digital image and the compressed image and then again using the decompressed image. In the first study compressed file sizes ranged from 5 to 50 K. It was found that for the larger file sizes both coding techniques yielded similar results, but for the lower file sizes, fractal transform coding yielded significantly better quality imagery than the DCT. In the second study compressed file sizes ranged from 6 to 90 kbytes. It was found that the DCT broke down for compression ratios greater than 200, while this only occurred for compression ratios well in excess of 800 for fractal transform coding. For all compression ratios considered in the second study, it was found that the rms error differences for the DCT were always higher, i.e. of lower fidelity, than the corresponding images from fractal compression. This was borne out by visual inspection of the resulting images in which the JPEG images displayed block artifacts as the limit of the JPEG technique was

approached. Thus, Sloan concludes that the fractal transform coding permits much smaller file sizes to be attained.

A novel idea put forward by Zhao and Yuan [51] is to combine fractal transform coding with the DCT. They have claimed that although fractal transform coding can achieve high compression, the quality of the decompressed image is not good. Their new method partitions the original image into 8X8 range blocks and 16X16 domain blocks denoted by $F_R(u,v)$ and $F_D(u,v)$ respectively. After these are transformed by the DCT, the range blocks are classified according to their AC coefficients into simple and complicated range blocks. A simple range block is approximated by storing its DC coefficient $F_R(0,0)$ which only requires 10 bits. A complicated range block is approximated by,

$$F_R(u,v) = \tau \circ \varphi(F_D(u,v)), \tag{8}$$

where, φ is a contractivity operator mapping domain blocks onto range blocks and τ is a compound transformation consisting of one of eight possible isometries that have been modified into DCT forms, a scaling and a luminance shift. The method then searches for the best matching domain block. Altogether 27 bits are required to approximate a complicated range block, which include 10 for the coordinates of the best matching block, 3 for the scaling factor, 3 for the isometry and 11 for the luminance shift. Zhao and Yuan have applied their method to the *Lena* image achieving a compression ratio of only 12.4 but a very good SNR of 32.3 dB. However, there is no mention of encoding and decoding times.

Finally, a new technique [52] is currently being developed that incorporates the high compression capabilities of fractal transforms into a DCT based compression algorithm as recommended by JPEG and MPEG. The idea is to extract the high frequency information or features of an input image from the low frequency information. The extracted features are to be encoded into fractals while the remainder of the image is to be compressed by a combination of DCT, VQ and entropy coding following the MPEG specification. Since the features of an image will undergo fractal transform coding, it will mean a reduction in the number of required computations and thus, this technique offers the potential of achieving high compression rates and fast processing.

7. Wavelet Transform Coding

The term *wavelet* was introduced at the beginning of the 1980s by a French geophysicist, J. Morlet [53,54]. It denotes a univariate function $\psi \in \mathbf{R}$, which, when subjected to the fundamental operations of integer shifts and dyadic dilations yields an orthogonal basis of $L_2(\mathbf{R})$. Such a function is called an orthogonal wavelet which can be applied to a finite group of data. Functionally, it is very much like the Discrete Fourier Transform where the input signal is assumed to be a set of discrete time samples. A

wavelet can be viewed as a bump which can be squeezed or expanded by a dilation and shifted by a translation. Wavelet coefficients can be efficiently computed and functions reconstructed from these coefficients using algorithms known as the wavelet transforms [54]. For a mathematical description the reader is referred to Appendix C.

7.1 Image Compression using Wavelets

Wavelet techniques have attracted much interest over the last few years because they not only eliminate the distortion that arises from data blocking, they also bring about a reduction in the block artefacts associated with Fourier based spectral methods such as the DCT. Furthermore, they can be employed to take advantage of the piecewise polynomial nature of real world images [55]. In essence, wavelet techniques are able to condense a large percentage of the total image into low frequency terms and can be used to approximate functions with little smoothness, a particularly useful feature with regard to image compression [54].

7.2 Implementation

The major deficiency of wavelet reconstruction is that the deepest nested dilation from decomposition must be the first to be reconstructed. This means that transformed data must be saved in memory so that the output appears in the reverse order in which it is calculated. Thus, the size of the input blocks and resolution in wavelet decomposition are limited by available memory. In fact, most of the effort in wavelet transform coding is in scheduling the filters and managing the input and output.

Hoag and Ingle [56] used the pyramid approach to wavelet data compression with vector quantisation as opposed to the commonly used scalar quantisation in which only the most significant bits of the wavelet coefficients are kept. Their aim was to compress underwater video data onboard an Autonomous Underwater Vehicle or AUV to enable it to be transmitted acoustically to a remote site. To support this application, the data must be massively compressed. They used a 256×256 test image from a clip of underwater video taken of the Titanic. The test image exhibited low contrast and detail inherent in underwater imagery. The best results were obtained with a 5-step wavelet decomposition in which the higher subbands coefficients were set to zero. The reconstruction of the Titanic image yielded a PSNR of 31.7 dB and the quality was excellent. Hoag and Ingle then made a comparison with the JPEG DCT algorithm. They found that in the low bit-rate (high compression) range between 0.1 and 0.2 bits/sample, the quality of the JPEG reconstructed images dropped off dramatically due to the inherent *blockiness* distortion caused by zeroing too many of the high frequency DCT coefficients. The Wavelet/VQ approach achieved much better PSNR results at the low-bit rates while better quality reconstructed images were obtained with the JPEG algorithm for bit rates greater than 0.2 bpp. In particular, Hoag and Ingle found that at 0.16 bpp, or a compression ratio of 50:1, the Wavelet/VQ

approach produced a degraded image due to blurring that remained intelligible whereas the JPEG image was distorted beyond recognition.

Zettler et al [55] were able to compress *Lena* images to ratios of 100:1 (0.08 bpp) and 50:1 (0.16 bpp) using wavelet transform coding. At 100:1 the decompressed image was distorted. Despite the high noise level, however, the features of the first reconstructed image were preserved as well as edges and general shapes. The distortion was restricted primarily to textures. For example, a *halo* effect was produced in the region immediately surrounding the image. Zettler et al also claim that advanced techniques can be employed to reduce the apparent distortion in images. The second decompressed image possessed considerably more fidelity. Zettler et al also concluded that with a custom chip implementation, entire multiplication lookup tables can be pre-loaded so that performance can be markedly improved by reducing the time required to carry out computation, which is necessary if a real-time capability is to be achieved with discrete wavelet transform (DWT) coding. It should also be mentioned that Zhang Ye et al [57] have also employed VQ in conjunction with DWT coding on 256X256X8 bit images and have obtained SNRs of 26.04 and 23.06 dB at coding rates of 0.78 and 0.70 bpp respectively.

In an interesting approach Rinaldo and Calvagno [58] have combined fractal transform coding with wavelet transform coding. First, the original images undergo wavelet decomposition whereupon each subimage is divided into range blocks. The range blocks are then matched with domain blocks chosen in the four lowest resolution subimages and coded through a description of the map that transforms the domain block into a range block. Rather than recursively coding range blocks from the blocks in the image, Rinaldo and Calvagno predict the range blocks of the subimage from the blocks of low resolution images which they claim simplifies the decoding procedure considerably and allows a more accurate control of the reconstruction error. This image decomposition technique acts as an automatic classifier of blocks, thereby reducing the block searching time and yielding smaller mean squared errors. As a consequence, Rinaldo and Calvagno state that their Wavelet-Fractal Coder (WFC) provides an improvement in both the compression rate and computational time.

Rinaldo and Calvagno applied their WFC to a 512×512X8 grey-level image of the *Lena* image. First, they present the original and reconstructed image at 0.25 bpp (32:1). The visual quality of the reconstructed image was fairly good, but some artefacts and ringing effects were noticeable near the edges. They also found that the WFC performed better than JPEG coding over an entire range of bit rates yielding an improvement in PSNR that was almost independent of bit rate. The total coding time for the image was about 2 mins on a Sun SPARC workstation with similar times involved for other images. Thus, the coding time is slightly longer than that by JPEG, which brings into question its suitability for real-time applications at present.

8. Discussion

So far, we have reviewed the current state of the art on three of the most prominent image compression techniques, but have not made any comparison between them which we aim to carry out in this section. Our comparison will be hampered by the shortage of literature directly comparing the techniques and the fact that we have been unable to analyse existing software employing the techniques.

Before proceeding any further we shall be required to give an indication of what we consider to be an acceptable image quality. Although our choice may be open to debate, we are going to adopt the *rule of thumb* that an image compression algorithm yielding a PSNR of more than 30 dB is acceptable.

Fisher et al [59] have made a comparison of fractal transform coding with the EPIC (Efficient Pyramid Image Coder) wavelet compression routine and JPEG compression. Their results are preliminary since the encoding time was not considered as a factor and that none of the codes had been adequately optimised. Because the degree of optimisation varies greatly for each technique, it may overshadow the strength of a particular compression technique. When encoding in fractals Fisher et al considered both quadtree and HV-rectangular partitioning approaches.

The images used in their study were 512X512X8 bit versions of *Lena* and the *Boat* benchmark image. The results are markedly different for PSNRs greater than 30 dB compared with those lower than 30 dB. For lower PSNRs JPEG compression yields a much lower compression ratio than the other approaches. For PSNRs greater than 30 dB, fractal transform coding employing quadtree partitioning yields the lowest compression ratio. For PSNRs greater than 35 the three remaining techniques yield almost identical compression ratios while for PSNRs between 30 and 35 there is a noticeable difference with HV fractal encoding offering the highest compression ratios. Specifically, for a PSNR of 30 dB JPEG coding offers a compression ratio of about 35:1, whereas the EPIC wavelet software and the fractal encoder with HV-rectangular partitioning offer ratios of about 45:1 and 55:1 respectively.

Fisher et al also present decoded images of the Boat benchmark image. They present the original image first and then give the JPEG coded version at a compression ratio of 54:1 (0.147 bpp). Here the PSNR is 23.7 dB and it is quite clear that many of the distinctive features in the original image such as the lighthouse and parts of adjacent boats have become severely degraded. The EPIC Wavelet version is presented at a compression ratio of 58:1 (0.138 bpp) and a PSNR of 26.4 dB. The image is a much better quality image than the JPEG image, but is not as fine as the fractal version, which has a compression ratio of 58.1:1 and a PSNR of 27.2 dB. Although the fractal version is the best of the images, very fine detail such as the boat's name is not as conspicuous as in the original image.

We have already mentioned that the reception of 320X200X8 images from the p-UAV at a rate of 1 frame/s requires a transmission rate of about 0.6 Mbits/s. For transmission in the VHF and higher frequency bands this does not present a problem, but unfortunately, it does mean that the p-UAV system can only be deployed in LOS operations. For non-LOS operations trans-mission is possible in the HF range provided permission can be obtained to combine four pre-allocated bandwidths as described in Section 3, thereby extending the bandwidth to 10 kbits. Then the data need only be compressed by a factor of 64. We have seen that this compression ratio is almost achievable with fractal compression, but not with wavelet and DCT encoding. Of the remaining two techniques the EPIC wavelet routine offers significantly better compression ratios. However, if the frame rate were reduced to 0.5 Hz or slightly lower, then it would be possible to transmit images by employing DCT coding. In addition, although a compression ratio of 64:1 is almost achievable with fractal encoding, the problem with fractal encoding is whether the encoding can be accomplished sufficiently quickly to meet near real-time requirements. Of course, this is one of the major topics in fractal transform coding currently under investigation as discussed in Section 6.

From our discussion of the three data compression techniques it can be seen that the DCT is the technique offering the closest to a near realtime capability. Rinaldo and Calvagno [58] mention that their WFC takes slightly longer than the DCT while fractal encoding takes significantly longer than the other two methods. Furthermore, to achieve higher compression ratios more processing time is required. Thus, parallel implementation, software optimisation and improvements in processor technology are still required before the other techniques will be able to match the current processing speed of the DCT.

Transmission of 640X400X24 (VGA quality) images at a TV frame rate of 25 Hz as described under naval operations in Section 3 requires a transmission rate of at least 144 Mbits/s, which means that transmission can only take place in the VHF or UHF bands and only after significant compression (greater than 30:1) has been applied. For these operations the range of the vehicle is, therefore, limited to LOS applications. In addition to the LOS limitation, the range is dependent on the signal strength and the gains in the antennas for the links.

Because of its compactness, the vehicle would possess a limited power and operate with an omnidirectional antenna. However, the GCS would be able to have a high gain directional antenna capable of high transmission power, thereby allowing the range to be extended. Typical UAV data link ranges with this arrangement range between 40 and 50 km. For the p-UAV the available power at the GCS will be certainly less than typical UAVs and the antenna pointing accuracy less precise. This means that a larger beam width would be required to monitor the p-UAV resulting in a lower maximum range.

In order to reduce the size of the bandwidth further after compression has been applied to the VGA images, the only remaining option is to reduce the frame rate since

we have seen that compression ratios at best range from 30:1 for JPEG compression to 55:1 for fractal encoding for PSNR values of 30 dB. Although this means transmission would not appear to be continuous to operator at the GCS, operating with a smaller bandwidth opens up the following possibilities;

(a) Transmission at the original frequency and power leads to better Signal to Noise Ratios, or in the case of digital data reduced Bit Error Rates (BER). This improvement can be used to extend the range or provide better immunity to external interference.

(b) The transmission power may be reduced by trading the improved SNR/BER against transmission power.

(c) The improved SNR/BER could also be traded for wider antenna beam width (thereby reducing antenna pointing requirements) or longer range.

9. Conclusion

In summary, the downlink for the p-UAV system in the short term could be based on DCT compression operating in the VHF/UHF bands and would have an LOS range of about 30 km using a directional antenna. For the naval operations mentioned in Section 3, however, transmission of VGA quality images would almost certainly require a reduced frame rate from the TV rate of 25 Hz. For land based operations where the resolution is not as critical and hence, can be reduced significantly, it would be possible to transmit images at the TV frame rate. Alternatively, less compressed or better quality images could be transmitted at the reduced frame rate.

A longer term aim would be to employ fractal and/or wavelet transform coding techniques to carry out the image compression. With continuing research into these data compression techniques and further advances in microprocessor technology, there is more than a real possibility that these techniques will offer a near realtime capability with compression ratios significantly higher than DCT compression in the not too distant future.

10. References

1. Cameron, K. and Kowalenko, V. (1995) *Portable Unmanned Aircraft System Concept Investigation,* DSTO-TR-0210, Melbourne.

2. Hartz, W.G., Alexovich, R.E. and Neustadter, M.S. (1989) *Data Compression Techniques Applied to High Resolution High Frame Rate Video Technology,* NASA Contractor Report 4263.

3. Langley, G. (1983) *Telecommunications Primer,* Pitman Books Ltd, London.

4. Schwartz, M. (1970) *Information Transmission, Modulation, and Noise,* McGraw-Hill Kogakusha, Tokyo.

5. Gardiner, P. (1989) *Electronics World + Wireless World,* Sept, pp. 920-923.

6. Sharman, J.R. (1988) *RPVs Remotely Piloted Vehicles,* Seventh International Conference, Bristol UK, 12-14 Sept, p. 21.1.

7. Hullett, J.L. and Budrikis, Z.L. (1972) *Proc. IREE* **33**, No. 11, pp. 518-524.

8. Frost, M.D., Burgess, M.E. and Thomson, P.L. (1977) *RANRL Report No.* 1/76 AR 10005, Vol.1, R.A.N. Research Laboratory, Edgecliff, NSW, (Secret).

9. Rogerson, S. (1991) *Electronics World + Wireless World,* pp. 30-33.

10. Davey, R.F. (1989) *Divisional Note* CSF/G12, Combat Systems Division, WSRL, DSTO Salisbury, S.A., Australia (Restricted).

11. Kayton, M. and Fried, W.R. (Eds.) (1969) *Avionics Navigation Systems* John Wiley & Sons, N.Y., p. 7.

12. Maslin, N. (1987) *HF Communications - A Systems Approach,* Pitman, Bath.

13. AGARD Lecture Series (1982) *Electromagnetic Propagation Problems in the Tactical Environment,* AGARD-LS-120.

14. Lathi, B.P. (1965) *Signals, Systems and Communication,* John Wiley & Sons, N.Y.

15. Walters, L.C. (1989) *Electronics World + Wireless World,* pp. 48-51.

16. Fiedler, D.M. and Hagn, G.H., (1983) *Army Communicator Fall,* pp. 14-22.

17. Fiedler, D.M. (1986) *Army Communicator Spring,* pp. 41-45.

18. Brune, J.F. and Ricciardi, B.V. (1979) *Modern HF Communications for Low Flying Aircraft*. This paper appears in AGARD Spec. Topics in HF Propagation (N80-19732 10-32).

19. Hooton, E.R. and Munson, K. (Eds.) (1995) *Jane's Battlefield Surveillance* 1994-1995, Jane's Information Group, Surrey, p. 236.

20. Beazley, K.C. (1987) *The Defence of Australia*, Aust. Govt. Publ. Service, Canberra, p. 55.

21. Hullett, J.L. (1972) *Proc. IREE* **33**, pp. 263-268.

22. Johns, P.B. and Rowbotham, T.R. (1972) *Communication Systems Analysis*, Butterworths, London.

23. Schapel, J.G. (1974) *TCS Group Working Paper No.* 31, DSTO Salisbury, S.A. (Restricted).

24. Rejman, M.H. (1988) *RPVs Remotely Piloted Vehicles* Seventh International Conference, Bristol UK 12-14 Sept. 1988, Supplementary Papers, p. 22.1.

25. Ahmed, N., Natarjan, T. and Rao, K.R. (1974) *Discrete Cosine Transform*, IEEE Trans. on Computers **C23**, pp. 90-93.

26. Barnsley, M.F. and Hurd, L.P. (1993) *Fractal Image Compression*, AK Peters Ltd, Wellesley, Massachusetts.

27. Khataie, M. and Soleymani, M.R. (1994) *Two Stage Residual Lattice Based Vector Quantisation*, Proceedings of the 1994 Canadian Conference on Electrical and Computer Engineering **1**, pp. 268-271.

28. Walmsley, N.P., Skodras, A.N. and Curtis, K.M. (1994) IEEE Trans. on Consumer Electronics **40**, pp. 11-19.

29. Eude, T., Cherifi, H. and Grisel, R. (1994) *Statistical distribution of DCT coefficients and their application to an adaptive compression algorithm*, Proceedings of TENCON'94 **1**, pp. 427-430.

30. Hung, A.C. and Meng, T.H.Y. (1994) *Statistical Inverse Discrete Cosine Transforms for Image Compression*, SPIE Vol. 2187, pp. 196-207.

31. McMillan, L. and Westover, L. (1992) *A forward-mapping realization of the inverse discrete cosine transform*, Data Compression Conference 1992, pp. 219-228.

32. Jung, S.H. and Mitra, S.K. (1994) *Improved DCT-based Image Coding and Decoding Methods for Low-Bit Rate Applications,* SPIE **2186** Image and Video Compression, pp. 156-162.

33. Loeffler C., Ligtenberg, A. and Moschytz, G.S. (1989) *Practical fast 1-D DCT algorithms with 11 multiplications,* ICASSP 89, pp. 988-991.

34. Linde, Y., Buzo, A. and Gray, R.M. (1980) *An algorithm for vector quantizer design,* IEEE Trans. on Communications **28**, pp. 84-95.

35. Wang, C.L. and Chen, C.Y. (1995) *High Throughput VLSI Architectures for the 1-D and 2-D Discrete Cosine Transforms,* IEEE Trans. on Circuits and Systems **5**, pp. 31-40.

36. Kung, H.T. and Leiserson, C.E. (1979) *Systolic arrays (for VLSI)* in Sparse Matrix Proc. 1978, I.S. Duff and G.W. Stewart (eds), Knoxville, TN: Society for Industrial and Applied Mathematics, 1979, pp. 256-282.

37. Mariatos, E.P., Metafas, D.E., Hallas, J.A. and Goutis, C.E. (1994) *A Fast DCT Processor, Based on Special Purpose CORDIC Rotators,* Proceedings of the 1994 IEEE International Symposium on Circuits and Systems **4**, pp. 271-274.

38. Fisher, Y. (1995) (Ed.) *Fractal Image Compression,* Springer-Verlag, New York.

39. Jacquin, A. (1993) *Fractal Image Coding: A Review,* Proceedings of the IEEE **81**, pp. 1451-1465.

40. Sloan, A.D. (1994) *Low Bit Rate Fractal Image Coding,* SPIE **2339** Visual Processing III, pp. 211-213.

41. Bani-Eqbal, B. (1995) *Speeding up Fractal Image Compression,* IEE Colloquium 'Fractals in Signal and Image Processing' Digest No.1995/006, pp. 2/1-2/3.

42. Fisher, Y., Jacobs, E.W. and Boss, R.D. (1990) *Fractal image compression using iterated transforms* in *Image and Text Compression,* Kluwer Academic Publishers, Boston, pp. 35-61.

43. Jacobs, E.W., Fisher Y. and Boss, R.D. (1992) *Image compression: A study of the iterated transform method,* Signal Processing **29**, pp. 251-263.

44. Popescu, D.C. and Yan, H. (1994) *Adaptive Fractal Image Compression,* International Symposium of Information Theory and its Applications, Sydney, Australia, 20-24 Nov. 1994, pp. 155-159.

45. Monro, D.M. and Dudbridge F. (1992) *Fractal approximation of image blocks* Proceedings ICASSP, pp III:485-III:488.

46. Monro, D.M. (1993) *A Hybrid Fractal Transform,* Proceedings ICASSP, pp. V:169-172

47. Monro, D.M. and Woolley, S.J. (1994) *Fractal Image Compression Without Searching,* Proceedings ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing **5**, pp. 557-560.

48. Monro, D.M. and Dudbridge, F. (1995) *Rendering Algorithms for Deterministic Fractals,* IEEE Computer Graphics and Applications, pp. 32-41.

49. Xue, M., Hanson, T. and Mérigot, A. (1994) *A Massively Parallel Implementation of Fractal Image Compression,* Proceedings ICIP-94, pp. 640-644.

50. Sloan, A.D. (1993) *Fractal Image Compression: A Resolution Independent Representation for Imagery,* Space and Earth Science Data Compression Workshop, NASA Conference Publication 3183, pp. 73-79

51. Zhao, Y. and Yuan, B. (1994) *Image compression using fractals and discrete cosine transform,* Electronics Letters **30**, pp. 474-475.

52. Jianmin Jiang (1995) *Image Compression with Fractals,* IEE Colloquium 'Fractals in Signal and Image Processing' Digest No.1995/006, pp. 7/1-7/3.

53. Koornwinder, T. H. (ed.) (1993) *Wavelets: An Elementary Treatment of Theory and Applications,* World Scientific Publishing, Singapore.

54. DeVore, R.A. and Lucifer, B.J. (1995) *Wavelets* in *Acta Numerica 1992* Cambridge University Press, Cambridge, pp. 1-56.

55. Zettler, W.R., Huffman, J. and Linden, D.C.P. (1990) *Application of Compactly Supported Wavelets to Image Compression,* SPIE **1244** Image Processing Algorithms and Techniques, pp. 150-160.

56. Hoag, D.F. and Ingle, V.K. (1994) *Underwater Image Compression Using the Wavelet Transform,* 1994 IEEE Oceans Proceedings, pp. II-533-II-537.

57. Zhang Ye, Zhao Guotian and Yu Hengchun (1993) *Wavelet Transform and Its Application in Image Compression,* IEEE TENCON'93, pp. 418-421.

58. Rinaldo, R. and Calvagno, G. (1994) *An Improved Wavelet-Fractal Coder, Proceedings* - IEEE International Symposium on Circuits and Systems **3**, pp. 113-116.

59. Fisher, Y., Rogovin, D. and Shen, T.P. (1994) *A Comparison of Fractal Methods with DCT and Wavelets,* SPIE **2304**, pp. 132-143.

60. Rabiner, L.R. and Gold, B. (1975) *Theory and Application of Digital Signal Processing*, Prentice-Hall, Englewood Cliffs, New Jersey.

61. Makhoul, J. (1980) *A Fast Cosine Transform in One and Two Dimensions*, IEEE Transactions on Acoustics, Speech, and Signal Processing **ASSP-28** (1), pp.27-34.

62. Barnsley, M.F. (1988) *Fractals Everywhere*, Academic Press, San Diego.

63. Strang, G. (1989) *Wavelets and Dilation Equations: A Brief Introduction* SIAM Review **31**, pp. 614-627.

64. Daubechies, I. (1988) *Orthonormal bases of compactly supported wavelets*, Comm. Pure and Appl. Math. **41**, pp. 909-996.

65. Edwards, T. (1992) *Discrete Wavelet Transforms: Theory and Imple-mentation*, Stanford University Preprint, Ca, US.

66. Mallat, S. (1989) *A theory for multiresolution signal decomposition: the wavelet representation*, IEEE Trans. Pattern Anal. Machine Intell. **11**, pp. 674-692.

Appendix A

Mathematical Details of the Discrete Cosine Transform

As mentioned in Section 5, the DCT's success as an image compression technique lies in its ability to eliminate the less visually stimulating high frequency components of a signal and to retain the quantised values of the low-frequency Fourier coefficients. In this appendix we present the mathematical details that are necessary for understanding how an algorithm based on the DCT can be developed.

Signals are defined in terms of discrete values of the independent time variable and are represented mathematically as sequences of numbers. A discrete-time system is essentially an algorithm for converting one of these sequences (an input) to another (an output) [60]. If $x(n)$ represents an input sequence and $y(n)$ an output sequence, then the response $h(n)$ of a system to a digital impulse is defined as,

$$y(n) = \sum_{m=-\infty}^{\infty} h(m)x(n-m). \tag{A1}$$

Introducing $x(n)=exp(i\omega n)$, where ω is the frequency, into the above equation yields,

$$y(n) = e^{i\omega n} \sum_{m=-\infty}^{\infty} h(m)e^{-i\omega m} = x(n)H(e^{i\omega}), \tag{A2}$$

where, $H(e^{i\omega})$ is the Fourier Series representation of the impulse response [60].

The Discrete Fourier Transform (DFT) is obtained by considering a sequence $x(n)$ with period N such that,

$$x(n) = \sum_{k=-\infty}^{\infty} X(k)\exp(2\pi ik) \ , \tag{A3}$$

where $\omega = 2\pi k / N$ are the only possible frequencies. Because of the periodicity $x(n)$ can be simplified to,

$$x(n) = \sum_{k=0}^{N-1} X(k)\exp(2\pi ik), \tag{A4}$$

while the DFT becomes,

$$X(k) = \sum_{n=0}^{N-1} x(n) \exp(-2\pi ikn / N). \tag{A5}$$

Following Ref. [61] we consider the signal to be a 2N-point even extension of the discrete-time signal $x(n)$ so that,

$$g(n) = \begin{cases} x(n), & 0 \le n \le N-1 \\ x(2N-n-1), & N \le n \le 2N-1 \end{cases}. \tag{A6}$$

As a consequence, the DFT for $g(n)$ can be written as,

$$X(k) = \sum_{n=0}^{2N-1} g(n) W_{2N}^{nk}, \tag{A7}$$

where $W_{2N}^{nk} = \exp(-i\pi nk / N)$, while its inverse known as the IDFT is given by,

$$g(n) = \frac{1}{2N} \sum_{k=0}^{2N-1} X(k) W_{2N}^{-nk}. \tag{A8}$$

Substituting Eq. (A6) into Eq. (A7) yields after some algebra,

$$X(k) = 2W_{2N}^{-k/2} \sum_{n=0}^{N-1} x(n) \cos\left[\frac{\pi(2n+1)k}{2N}\right] = W_{2N}^{-k/2} C(k), \tag{A9}$$

with,

$$C(k) = 2\sum_{n=0}^{N-1} x(n) \cos\left[\frac{\pi(2n+1)k}{2N}\right], \tag{A10}$$

and $0 \le k \le 2N-1$. Eq. (A9) is known as the *One Dimensional Discrete Cosine Transform* (1D-DCT) of the discrete time signal and its inverse, the 1D-IDCT, is given by

$$x(n) = \frac{1}{N}\left[\frac{C(0)}{2} + \sum_{k=1}^{N-1} C(k) \cos\left[\frac{\pi(2n+1)k}{2N}\right]\right], \tag{A11}$$

where $0 \le n \le N-1$. It is as a result of the even symmetry of the signal that the cosine factors have appeared.

Image data represent two-dimensional signals and thus the preceding material must be extended before it can be utilised as a data compression technique. The 2D DCT is obtained by defining the signal as a (2Nx2N)-point even extension in which,

$$y(n_1,n_2)=\begin{cases} x(n_1,n_2), & 0\le n_1\le N_1-1, \quad 0\le n_2\le N_2-1 \\ y(2N_1-n_1-1,n_2), & N_1\le n_1\le 2N_1-1, \quad 0\le n_2\le N_2-1 \\ y(n_1,2N_2-n_2-1), & 0\le n_1\le N_1-1, \quad N_2\le n_2\le 2N_2-1 \\ y(2N_1-n_1-1,2N_2-n_2-1), & N_1\le n_1\le 2N_1-1, \quad N_2\le n_2\le 2N_2-1 \end{cases}$$

By using the above results the 2D Discrete Fourier Transform of $y(n_1,n_2)$ can be expressed as,

$$Y(k_1,k_2)=\sum_{n_1=0}^{2N_1-1}\sum_{n_2=0}^{2N_2-1} y(n_1,n_2)W_{2N_1}^{n_1k_1}W_{2N_2}^{n_2k_2}, \tag{A12}$$

while its inverse (the 2-D IDFT) is given by,

$$y(n_1,n_2)=\sum_{k_1=0}^{2N_1-1}\sum_{k_2=0}^{2N_2-1} Y(k_1,k_2)W_{2N_1}^{-n_1k_1}W_{2N_2}^{-n_2k_2}. \tag{A13}$$

Substituting $y(n_1,n_2)$ into Eq. (A8), one obtains after some algebra,

$$Y(k_1,k_2)=W_{2N_1}^{-k/2}W_{2N_2}^{-k_2/2}C(k_1,k_2), \tag{A14}$$

where,

$$C(k_1,k_2)=4\sum_{n_1=0}^{N_1-1}\sum_{n_2=0}^{N_2-1} x(n_1,n_2)\cos\left[\frac{\pi(2n_1+1)k_1}{2N_1}\right]\cos\left[\frac{\pi(2n_2+1)k_2}{2N_2}\right], \tag{A15}$$

and $0\le k_j\le N_j-1$, $(j=1,2)$. Eq. (A15) is known as the *2D Discrete Cosine Transform* (2-D DCT) of the sequence $x(n_1,n_2)$ and its inverse, the 2-D IDCT, is given by,

$$x(n_1,n_2)=\frac{1}{N_1N_2}\sum_{k_1=0}^{N_1-1}\sum_{k_2=0}^{N_2-1} C'(k_1,k_2)\cos\left[\frac{\pi(2n_1+1)k_1}{2N_1}\right]\cos\left[\frac{\pi(2n_2+1)k_2}{2N_2}\right], \tag{A16}$$

where,

$$C'(k_1,k_2)=\begin{cases} C(0,0)/4, & k_1=0,\ k_2=0 \\ C(k_1,0)/4, & k_1\ne 0,\ k_2=0 \\ C(0,k_2)/4 & k_1=0,\ k_2\ne 0 \\ C(k_1,k_2) & k_1\ne 0,\ k_2\ne 0 \end{cases}$$

In JPEG compression the pixel values from 8x8 blocks of the original image are first adjusted to centre them at zero. For example, if pixel data are in an 8-bit format, then 128 is subtracted from them so that the signal can be regarded as even. Then the DCT or Eq. (A14) is applied to the normalised pixel values so that transformed 8-bit data are stored as 11 bit signed integers. It is these signed integers that are quantised by dividing them by a quantisation coefficient and rounding off to the nearest integer. The quantisation co-efficients vary based on the fact that quantisation can be much more severe for higher frequencies than for lower frequencies because of the the human visual system's relative insensitivity to high frequencies. In addition, because JPEG compression transforms 8x8 blocks, there is no need to evaluate the cosine factors in Eq. (A14) repeatedly. Instead, a look-up table can be em-ployed. For more details regarding implementation the reader is referred to Refs. 25 and 26.

Appendix B

Fractal Transform Coding

Here we review the basic mathematical concepts underpinning the implementation of deterministic fractals in image compression. For a more comprehensive treatment of the subject the reader is referred to Refs. [26,38,62].

Barnsley [26,62] was first to propose the idea of fractal image compression in which real-life images could be modelled by deterministic fractals. Deterministic fractals represent the fixed points of a set of two-dimensional affine transformations. As a consequence, the mathematics of *Iterated Function Systems* (IFS) and *Recurrent Iterated Functions Systems* (RIFS) together with the Collage Theorem have been developed to provide the theoretical foundation of fractal image compression. We shall describe IFSs here, although it should be pointed out that to encode images in an automated approach, one must use piecewise affine contractive transformations, which make use of only the partial self-transformability of images. Fisher [38] refers to these as *Partitioned Iterated Function Systems* or PIFS. Because PIFSs allow not only the encoding of grey scale images, but also partition an image into pieces that can be transformed separately, they are able to encode many shapes that cannot be encoded by IFSs.

The basic building block of present fractal image compression systems is the affine transformation which for two dimensions is defined as a mapping $w:\Re^2 \to \Re^2$ where $w(x,y) = (ax+by+e, cx+dy+f)$ and $a,b,c,d,e,f \in \Re$. Such a transformation can be represented in matrix form as,

$$w(\underline{x}) = w\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix} = \mathbf{A}\underline{x} + \mathbf{T} \qquad \text{(B1)}$$

If we consider the one affine transformation of $f(x) = ax+b$, $\forall x \in \Re$ over the interval [0,1], then the new interval length becomes $|a|$. Thus, the transformation f rescales the interval by a factor a while the left endpoint of the interval is translated to b. When $|a|<1$, the affine transformation is said to be contractive. For higher dimensions we require the theory of metric spaces to define a contractive transformation.

Let (X,d) denote the complete metric space of digital images where d is a given metric or distortion measure and let μ_0 denote the original image to be encoded. The goal of iterated transformation theory is the construction of a contractive image transformation, τ, defined from the space (X,d) onto itself, for which μ_0 is an approximate fixed point. This is known as the *Inverse Problem*. The transformation, τ,

is referred to as the *fractal code* for μ_0 while μ_0 is said to be approximately *self-transformable* under τ.

We mentioned above that deterministic fractals represent the fixed points of sets of two-dimensional affine transformations. The fixed point of a transformation $f: X \to X$ on a metric space (X,d) is the point $x_f \in X$ such that $f(x_f) = x_f$. A transformation $f: X \to X$ on a metric space (X,d) is Lipschitz if there is a constant s (known as a Lipschitz factor) such that $d(f(x), f(y)) \leq s.d(x,y) \quad \forall x, y \in X$. When $s < 1$, f is said to be *contractive* or a *contraction mapping.*

We are now in a position to give one of the fundamental results of fractal image compression:

Theorem 1-The Contraction Mapping Theorem

Let $f: X \to X$ be a contraction mapping on a complete metric space (X,d). Then f possesses a unique point $x_f \in X$ such that for any point $x \in X$, the sequence $\left\{f^{\circ n}(x) : n = 0,1,2,\ldots\right\}$ converges to x_f, i.e.

$$\lim_{n\to\infty} f^{\circ n}(x) = x_f \quad \forall x \in X. \tag{B2}$$

The point x_f is called a fixed point of the mapping f. The proof of this theorem can be found in Refs. [38,62]. This theorem states that the fixed point of a transformation f will be the image one gets when the sequence $f(x_0)$, $f(f(x_0))$, $f(f(f(x_0)))$,..., is computed for *any* image x_0. That is, as long as the transformation is contractive in the space of images, it will have a unique fixed point that will then be some image.

In fractal image compression it is convenient to use the Hausdorff space $H(X)$ where one can study compact subsets of metric spaces. This means that by using $H(R^2)$ one can only concentrate on drawings, pictures and other *black on white* subsets of R^2. In addition, when using this space, another metric is required, which is known as the *Hausdorff distance or metric* $h(A,B)$. For a complete metric space (X,d) this metric for points A and B in $H(X)$ is given by $h(A,B) = d(A,B) \vee d(B,A)$ where $x \vee y$ represents the maximum of x and y. That is, if A is an element of the associated Hausdorff space $H(X)$, then,

$$A_d(\varepsilon) = \left\{x \middle| d(x,y) \leq \varepsilon \text{ for some } y \in A\right\},$$

which means that $A_d(\varepsilon)$ is the set of points of maximal distance ε from A. The Hausdorff distance between two elements A and B of $H(X)$ becomes ,

$$h_d(A,B) = \max\left\{ \inf\left\{ \varepsilon \,\middle|\, B \subset A_d(\varepsilon) \right\}, \inf\left\{ \varepsilon \,\middle|\, A \subset B_d(\varepsilon) \right\} \right\}.$$

An *Iterated Function System* consists of a complete metric space (X,d) together with a finite set of contraction mappings $w_n : X \to X$ with respective contractivity factors s_n for n=1,2,...,N. An IFS is denoted by $\{X; w_n, n = 1,2,\ldots,\mathrm{N}\}$ with its contractivity factor $s = \max\{s_n : n = 1,2,\ldots,\mathrm{N}\}$. IFSs, or their generalisations mentioned above, are the basic building blocks of fractal transform coding. The PIFSs presented in Ref. [38] possess not only the two spatial dimensions of IFSs but also a third dimension for the grey levels of an image. This means that Eq. (B1) has to be modified to include a z component with **A** now becoming a 3x3 matrix and **T** a three dimensional column vector. The third row and column of **A** consist of zeros except for the diagonal element which consists of a term s_i to control the contrast and the additional element o_i in **T** controls the brightness of the transformation.

The following theorem proposes methods for constructing the fixed point (or attractor) of an IFS. Let $\{\Re^2 : w_1, w_2, \cdots, w_N\}$ be an IFS and choose a compact set $\mathrm{A}_0 \subset \Re^2$. Then a sequence $\{\mathrm{A}_n : n = 0,1,2,\cdots\} \subset H(\Re^2)$ can be constructed. According to this theorem known as the IFS theorem, the sequence $\{\mathrm{A}_n\}$ converges to the attractor of the IFS in the Hausdorff metric. Thus, we have a procedure for calculating successive approximations to the fixed point of an IFS.

Theorem 2-The IFS Theorem

Let $\{X; w_n, n = 1,2,\ldots,\mathrm{N}\}$ be an IFS with contractivity factor s. Then the transformation $W : H(X) \to H(X)$ defined by,

$$W(B) = \bigcup_{n=1}^{N} w_n(B) \quad \forall B \in H(X)$$

is a contraction mapping on the complete metric space $(H(X), h(d))$ with contractivity factor s. That is, $h(W(B), W(C)) \le s.h(B,C)$, $\forall B, C \in H(X)$. Its unique fixed point, $A \in H(X)$, obeys,

$$A = W(A) = \bigcup_{n=1}^{N} w_n(A), \tag{B3}$$

and is given by,

$$A = \lim_{n\to\infty} w^{\circ n}(B) \quad \forall B \in H(X).$$

This result is proved in Ref. [62]. The fixed point is called the *attractor* of the IFS and as a consequence of its uniqueness, we are led to the following theorem:

Theorem 3-The Collage Theorem

Let (X,d) be a complete metric space. Let $L \in H(X)$ and $\varepsilon \geq 0$ be given. Choose an IFS with contractivity factor $0 \leq s < 1$, so that $h\left(L, \bigcup_{n=1}^{N} w_n(L)\right) \leq \varepsilon$. Then $h(L,A) \leq \varepsilon/(1-s)$ where A is the attractor of the IFS.

This theorem, which is also proved in Ref. [62], states that for a given set or image L an IFS, or a set of contractive transformations, can be found for which L is the attractor. That is, the union or collage of the images of L under the transformations is *close to* or looks like L. The degree to which two images *look alike* is measured by using the Hausdorff metric which in turn depends on the metric d.

To summarise the above, let (X,d) be a complete metric space and $\mu \in H(X)$ be any given image. Given a set of contractive transformations such that $T: H(X) \to H(X)$, we know from the Contraction Mapping Theorem that,

$$\lim_{n\to\infty} T^{\circ n}(\mu) = x_T \quad \forall \mu \in H(X). \tag{B4}$$

From Theorem 2,

$$\lim_{n\to\infty} T^{\circ n}(\mu) = T(\mu) = \mu, \ \forall \mu \in H(X).$$

Since an attractor is unique, we have $\mu = x_T$ and μ must be formed of transformed copies of itself. According to the Collage Theorem, minimising the distance between μ and $T(\mu)$ (the collage of the image) minimises the distance between the fixed point x_T and μ. In practice, it is not possible to find a T such that $\mu = T(\mu)$, but it is possible to find a μ_{approx} satisfying $T(\mu_{approx}) = \mu_{approx}$. That is,

$$\mu \approx \mu_{approx} = T(\mu_{approx}) \approx T(\mu) \Rightarrow \quad \mu \approx T(\mu).$$

Now we can say that a fractal image is constructed from a 'collage' of transformed copies of itself and is thus inherently self-similar.

Consider an image as a surface lying over a plane, defined by a function f(x,y), that returns a number between 0 and 1 at each position (x,y); the range 0 to 1 can represent grey levels from black to white [38]. The original image μ_{orig} becomes a function mapping the unit square into the real numbers, i.e. $I^2 = [0,1] \times [0,1] \rightarrow \Re$. The image is now split into non-overlapping domain blocks D_i and range blocks R_i where the union of the domain blocks yields the original image, i.e. $\bigcup_i D_i = I^2$. Furthermore, we map the domain blocks into range blocks by a collection of affine transformations so that $D_i \rightarrow T(D_i) = R_i$. This means that the fractal compression scheme has been reduced to a search through all the range blocks, viz. the set of all R_i's, to find an R_i for each D_i, which minimises some measure of distortion or similarity. That is, we search for the part of the image that most looks like the part of the image in the domain block [38]. Specifically, we seek a transformation $g \in T$ where,

$$\forall\ \mu, \upsilon \in X,\ \exists\ s < 1,\ \text{ such that } d(g(\mu), g(\nu)) \leq d(\mu, \nu) \text{ and } d(\mu_{orig}, g(\mu_{orig})),$$

is as 'close to zero' as possible. By repeated use of the triangle inequality, it can be shown for any image μ_0 and any positive integer n that,

$$d(\mu_{orig}, g^n(\mu_0)) \leq (1-s)^{-1} d(\mu_{orig}, g(\mu_{orig})) + s^n d(\mu_{orig}, \mu_0).$$

From this result it can be seen that, after a number of iterations, the terms of any iterated sequence of the form $\{\mu_n = g^n(\mu_0)\}_{n \geq 0}$, where μ_0 is some arbitrary initial image, cluster around the original image. In a space of discrete images, the sequence converges exactly to a stable image which is its attractor. The closeness of $g^n(\mu_0)$ to μ_{orig} is determined by the measure or distortion $d(\mu_{orig}, g(\mu_{orig}))$, which is generally taken to be the *root mean square* difference between image blocks as described in Section 4 and which is known in mathematical terms as an L^2 metric.

A fractal code is obtained from the search and represents a statement such as 'region A of an image is most like region B after transformation'. The original image or rather the pixel data are not transmitted to a decoder, only the fractal code. Thus, decoding an image consists of repeatedly applying the transformations in the fractal code to an arbitrary initial image I_0 until the images converge to a fixed point [38].

Appendix C

Wavelet Transform Coding

Wavelets represent a family of basis functions derived from one single function subjected to shifts and dilations. In this appendix we present the basic properties of wavelets, which are necessary for understanding and developing a wavelet transform coding scheme in image compression.

It is shown in [63] that it is not necessary for the set of wavelet functions given by the recursive equation of $\psi_{j,k} := 2^{k/2}\psi(2^k \cdot - j)$ to form a complete orthonormal set. However, when they do any function, $f \in L_2(R)$, can be decomposed into a series of the form,

$$f = \sum_{j,k \in Z} \langle f, \psi_{j,k} \rangle \psi_{j,k} , \qquad \text{(C1)}$$

where $\langle f, g \rangle := \int f \bar{g} dx$ is the usual inner product of two $L_2(R)$ functions. Eq. (C1) can be viewed as the construction of f from bumps $\psi_{j,k}$ (functions with compact support) with small values of k contributing to the broad resolution of f and large values of k producing the finer detail. The decomposition given by Eq. (C1) is analogous to a Fourier decomposition of f in terms of the exponential functions $e_k := e^{ik\cdot}$, although important differences exist. For example, all terms in the Fourier series contribute to the value of f at a point x while wavelets are usually of compact support or fall off exponentially at infinity. Thus, the only terms in Eq. (C1) corresponding to $\psi_{j,k}$ with $j \cdot 2^{-k}$ near x yield a large contribution at x. Hence, the representation can be regarded as local.

Multiresolution analysis is a method of creating an orthonormal wavelet basis by breaking $L_2(R)$ up into a sequence of closed subspaces V_j in the form of

$$\cdots \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \subset \cdots , \qquad \text{(C2)}$$

where $V_m \rightarrow L_2(R)$ as $m \rightarrow \infty$. These subspaces are subject to the following properties;

(i) $V_n \subset V_{n-1}$, $n \in Z$,
(ii) $\bigcup_{n=-\infty}^{\infty} V_n$ is dense in $L_2(R)$ and $\bigcap_{n=-\infty}^{\infty} V_n = \{0\}$,
(iii) $f(x) \in V_n \Leftrightarrow f(2x) \in V_{n-1}$,
(iv) $f(x) \in V_0 \Leftrightarrow f(x-k) \in V_0 \ \forall k \in Z$, and
(v) $\exists\, g \in V_0$ such that $g(-k)$, $k \in Z$ is a Riesz basis for V_0.

A *Riesz* basis is a set $\{x_n\}$ in a Hilbert space H where an orthonormal basis $\{e_n\}$ and a bounded linear operator T are related by

$$Te_n = x_n. \qquad , \forall \text{ n.} \tag{C3}$$

In addition to the above, there exists the following rule concerning the speed of oscillations,

$$f \in V_m \Leftrightarrow f(2\cdot) \in V_{m-1},\ m \in Z. \tag{C4}$$

So if f is an oscillating function in V_m, then the function oscillates twice as fast as an element of V_{m-1}.

In wavelet theory it is assumed that V_0 is generated by the integer translates $\phi_{0n}(x)=\phi(x-n)$ of one single function ϕ known as the father. Each $f \in V_0$ can be written as,

$$f = \sum_{n=-\infty}^{\infty} a_n \phi_{0n}\,. \tag{C5}$$

Since $\phi \in V_0$ and $V_0 \subset V_{-1}$, we have $\phi \in V_{-1}$ from Eq. (C4) and $\phi(2^{-1}\cdot) \in V_0$. Thus, the wavelet basis is given by the recursive difference equation,

$$\phi(x) = \sum_{n=-\infty}^{\infty} c_n \phi(2x-n),\ \ x \in \Re, \tag{C6}$$

for some coefficients $\{c_n\}$. The range of the summation in Eq. (C6) is determined by the number of nonzero coefficients, is arbitrary and is also referred to as the order of the wavelet. Rearranging Eq. (C6) yields,

$$\phi(x) = \sqrt{2} \sum_{-\infty}^{\infty} h_n \phi(2x-n),\ \ x \in \Re, \tag{C7}$$

where the factor of $\sqrt{2}$ arises from normalisation. The numbers $h_n=c_n/\sqrt{2}$ are called the filter coefficients of ϕ and obey the following condition,

$$\sum_{n=-\infty}^{\infty} h_n = \sqrt{2}\,. \tag{C8}$$

Eq. (C6) is orthogonal to its translations, i.e. $\int \phi(x)\phi(x-k)dx = 0$. We also desire an equation which is orthogonal to its dilations, i.e. $\int \varphi(x)\varphi(2x-k)dx = 0$. This is the associated wavelet or mother of the wavelets and is generated from ϕ by the following equation,

$$\psi(x) = \sqrt{2}\sum g_n \phi(2x-n),\ g_n = (-1)^n \overline{h_{1-n}}. \tag{C9}$$

From this, other related functions can be defined,

$$\phi_{m,n}(x) = 2^{-m/2}\phi(2^{-m}x-n) \quad m,n \in Z, \tag{C10}$$

with the corresponding wavelets given by,

$$\psi_{m,n}(x) = 2^{-m/2}\phi(2^{-m}x-n) \ \ m,n \in Z. \tag{C11}$$

The system $\{\psi_{n,k} \mid k,n \in Z\}$ is also called an *orthonormal wavelet* basis. In most applications the sums given above are finite and we consider this to be the case from here on.

Eq. (C6) can be solved by contructing an MXM matrix of coefficients, where M is the number of nonzero coefficients. This matrix can be designated by L with entries $L_{ij}=c_{2i-j}$. It always has an eigenvalue equal to unity and the respective normalised eigenvector consists of the value of ϕ at integer values of x. Once these values are known, all other values of ϕ can be generated by applying the recursion equation to get the values at half integers, quarter integers and so on to the desired dilation.

Plots of most wavelet functions appear to be extremely irregular, which is due to the fact that a wavelet function is non-differentiable everywhere. The functions that are normally used in transforms consist of a few sets of well-chosen coefficients which results in a function that has a discernible shape such as the Haar basis function or the Daubechies-4 wavelet. The latter is often used in data compression.

In applying wavelet theory to image compression, we note that pixel values can be predicted by considering the complete image as a histogram and then looking at the values of neighbouring pixels. Thus spatial correlations occurring in natural images are taken into account. To create a good image decomposition scheme based on this approach, the image is split into a low resolution part consisting of a smaller number of samples than the original image and a difference signal which is the difference between the low resolution part and the actual image. The low resolution part is actually a good estimate of the true image due to the correlations present in real world images. The image it generates will still contain spatial correlations and hence further

decompositions can take place, thereby creating a hierarchical decomposition of the original image.

An efficient decomposition scheme can be created by employing multi-resolution wavelet bases, which we describe here only briefly. A more detailed discussion appears in Refs. [64,65]. We let ϕ be the generator of a multiresolution wavelet basis, so that we can put,

$$\phi_{m,n}(x) = 2^{-m/2}\phi(2^{-m}x - n). \tag{C12}$$

The spaces V_m=*span*$\{\phi_{m,n} \mid n \in \mathbf{Z}\}$ correspond with different resolution levels of our decomposition. Then there is a function, ψ, such that the orthogonal space W_m=*span*$\{\psi_{m,n} \mid n \in \mathbf{Z}\}$ satisfies the direct sum given by,

$$V_{m-1} = V_m \oplus W_m, \tag{C13}$$

where,

$$\psi_{m,n}(x) = 2^{-m/2}\psi(2^{-m}x - n).$$

Furthermore, let P_m and Q_m represent the orthogonal projections on V_m and W_m respectively while the sequence $(c_n)_{n\in\mathbf{Z}}$ represents the signal undergoing compression. The projections P_n and Q_n are defined respectively as,

$$\begin{aligned} P_n f &= \sum_{n=-\infty}^{\infty} c_k^n(f)\phi_{n,k}, \text{ and} \\ Q_n f &= \sum_{n=-\infty} d_k^n(f)\psi_{n,k}. \end{aligned} \tag{C14}$$

The coefficients are the inner products shown below,

$$c_k^n(f) = < f, \phi_{n,k} > \text{ and } d_k^n(f) = < f, \psi_{n,k} >.$$

We define a sequence $(c_n^o)_{n\in\mathbf{Z}}$ with $c_n^o = c_n$ and an associated function by,

$$f(x) = \sum_n c_n^o \phi_{on}(x). \tag{C15}$$

Applying a multiresolution analysis to f means that f can be put equal to $P_1 f + Q_1 f$ where the first term is the low resolution representation of f and the second term represents the difference signal. Specifically,

$$P_1 f = \sum_k c_k^1 \phi_{1k}, \text{ and}$$
$$Q_1 f = \sum_k d_k^1 \psi_{1k}. \qquad \text{(C16)}$$

The coefficients c_k^1 are given by,

$$c_k^1 = < P_1 f, \phi_{1k} > = < f, \phi_{1k} >$$
$$= \sum_n c_n^0 < \phi_{on}, \phi_{1k} > = \sum_n c_n^0 h_{n-2k} , \qquad \text{(C17)}$$

where,

$$h_n = 2^{-1/2} \int \phi\left(x/2\right) \phi(x-n)\, dx.$$

The coefficients d_k^1 are evaluated in a similar fashion. By repeating the procedure N times we arrive at the decomposition,

$$f = Q_1 f + Q_2 f + \cdots + Q_N f + P_N f , \qquad \text{(C18)}$$

where,

$$P_n f = \sum_k c_k^n \phi_{nk} ,$$

and,

$$Q_n f = \sum_k d_k^n \psi_{nk} .$$

The coefficients c_k^n and d_k^n are determined from the following recursive relations,

$$c^j = H c^{j-1} \text{ and } d^j = G d^{j-1} , \qquad \text{(C19)}$$

where,

$$(\mathbf{H}a)_k = \sum_n h_{n-2k} a_n \text{ and } (\mathbf{G}a)_k = \sum_n g_{n-2k} a_n ..$$

After a number of iterations, the original image sequence c^o is decomposed into the lowest resolution signal c^N and the difference signals $d^N, d^{N-1}, \ldots, d^1$ of ever finer resolution. The above analysis is a one-dimensional multiresolution representation and can be extended to two dimensions by using products as opposed to sums in the above results. The reader is referred to pp. 86-87 of Ref. 53 for this non-trivial exercise.

In summary, the Discrete Wavelet Transform (DWT) in one dimension produces two output sequences, referred to as "odd" and "even", from an input sequence. These can be viewed as a pair of convolution functions or Finite Impulse Response (FIR) filters

[55]. Both filters create an output stream that is half the length of the original input. In many situations, the low-pass filter output or odd output contains most of the information content of the original signal and is related to Eq. (C6) by,

$$a_i = \frac{1}{2}\sum_{j=1}^{N} c_{2i-j+1} f_j \ , \quad i = 1,2,\ldots,N/2, \tag{C20}$$

where; c_{2i-j+1} are the wavelet coefficients, f_j is the input function of block size N and a_i are the odd output values. For the Haar wavelet there are only two coefficients, $c_0=1$ and $c_1=1$ while for the Daubechies 4-wavelet, there are four coefficients, $c_0=(1+\sqrt{3})/4$, $c_1=(3+\sqrt{3})/4$, $c_2=(3-\sqrt{3})/4$ and $c_3=(1-\sqrt{3})/4$. In general, higher order wavelets, i.e. those with more nonzero corefficients tend to put more information in the odd output and less into the even output.

The high-pass filter output or even output contains the difference between the true output and the value of the reconstructed input if it were to be reconstructed from only the information given in the odd output. The even output values b_i can also be expressed in terms of wavelet coefficients as,

$$b_i = \frac{1}{2}\sum (-1)^{j-1} c_{j+2-2i} f_j \ , \ i = 1,2,\ldots,N/2. \tag{C21}$$

An important step in wavelet data compression is determining those wavelet functions which result in the even terms being almost zero [53]. In fact, if the average amplitude of the even output is sufficiently low, then the even half of the signal can be discarded without significant degradation occurring in the reconstructed signal. Since most of the information is held in the low-pass filter output, this can again be transformed into two new sets of data. If the number of input samples is $N=2^D$, then a maximum of D dilations can be performed with the last dilation resulting in a single low-pass value and a single high-pass value. Thus successive dilations represent lower and lower frequency content by halves. In addition, to obtain high compression rates, it may be necessary to begin with large blocks of input so that not only more dilations can be carried out, but also lower frequencies can be represented in the decomposition.

Basically three parameters are required in implementing a wavelet coding scheme;

(i) the filter length reflecting the number of coefficients that describe the wavelet function,

(ii) the block size N of the input, which must be a power of two, and

(iii) the number of dilations or passes of the input stream.

For $N=2^D$, D dilations are possible for full decomposition, but this is not always suitable in compressing data [66].

After the transformation steps are completed quantisation is usually performed. A finite number of real-valued coefficients is selected to form a quantisation grid. Each coefficient is then replaced by the nearest point in the grid. The grid can be chosen by taking evenly spaced points or by choosing points closer together near zero. The quantised coefficients are then coded [53,55].

An exact reconstruction of the image can only be made if the decomposition coefficients are known exactly, which is not possible since they are not integers. Thus, quantisation of the coefficients is required, similarly as for the DCT discussed in Appendix A. In the one dimensional case the original image can be reconstructed by repeated use of the relation,

$$\begin{aligned} P_{j-1}f &= P_j f + Q_j f \\ &= \sum_n c_k^j \phi_{jk} + \sum_n d_k^j \psi_{jk}. \end{aligned} \tag{C22}$$

This implies that,

$$\begin{aligned} c_n^{j-1} &= < P_{j-1}f, \phi_{j-1,n} > \\ &= \sum_k c_k^j < \phi_{jk}, \phi_{j-1,n} > + \sum_k d_k^j < \psi_{jk}, \phi_{j-1,n} > \\ &= \sum_k c_k^j h_{n-2k} + \sum_k d_k^j g_{n-2k}. \end{aligned} \tag{C23}$$

Reconstruction of the original image in either one or two dimensions relies on the correct choice of the function ϕ.

By implementing an equally spaced quantisation scheme Nacken (p. 81 of Ref. [53]) has achieved bit rates as low as 0.4 bits per pixel (95% data reduction) whilst at the same time maintaining a high quality of reconstructed image. Fine details such as small bright spots are preserved better with this approach due to the localisation of the low level basis functions.

DISTRIBUTION LIST

Data Link Technology For A Portable Unmanned Aerial Vehicle

Victor Kowalenko, Jane Phipps and Keith Cameron

AUSTRALIA

DEFENCE ORGANISATION

S&T Program

Chief Defence Scientist
FAS Science Policy
AS Science Industry and External Relations
AS Science Corporate Management
} shared copy

Counsellor Defence Science, London (Doc Data Sheet)
Counsellor Defence Science, Washington (Doc Data Sheet)
Scientific Adviser to MRDC Thailand (Doc Data Sheet)
Director General Scientific Advisers and Trials/Scientific Adviser Policy and Command (shared copy)
Navy Scientific Adviser (3 copies Doc Data Sheet and 1 copy distribution list)
Scientific Adviser - Army (Doc Data Sheet and distribution list only)
Air Force Scientific Adviser
Director Trials

Aeronautical and Maritime Research Laboratory
Director

Electronics and Surveillance Research Laboratory
Director

Chief of Air Operations Division
Dr. V. Kowalenko (5 copies)
J. Phipps (1 copy)
K. Cameron (3 copies)

Information Technology Division
Dr. Murray Grigg (Canberra)
Dr. Garry Newsam
Dr. Janet Aisbett

DSTO Library
Library Fishermens Bend
Library Maribyrnong
Library DSTOS (2 copies)
Australian Archives
Library, MOD (Doc Data sheet only)

Forces Executive

Director General Force Development (Sea) (Doc Data Sheet only)
Director General Force Development (Land) (Doc Data Sheet only)
Director General Force Development (Air)

Army

ABCA Office, G-1-34, Russell Offices, Canberra (4 copies)

S&I Program

Defence Intelligence Organisation
Library, Defence Signals Directorate (Doc Data Sheet only)

B&M Program (libraries)

OIC TRS, Defence Central Library
Officer in Charge, Document Exchange Centre (DEC), 1 copy
*US Defence Technical Information Centre, 2 copies
*UK Defence Research Information Center, 2 copies
*Canada Defence Scientific Information Service, 1 copy
*NZ Defence Information Centre, 1 copy
National Library of Australia, 1 copy

UNIVERSITIES AND COLLEGES

Australian Defence Force Academy
Library
Head of Aerospace and Mechanical Engineering
Deakin University, Serials Section (M list), Deakin University Library, Geelong, 3217
Senior Librarian, Hargrave Library, Monash University
Librarian, Flinders University

OTHER ORGANISATIONS

NASA (Canberra)
AGPS

OUTSIDE AUSTRALIA

ABSTRACTING AND INFORMATION ORGANISATIONS

INSPEC: Acquisitions Section Institution of Electrical Engineers
Library, Chemical Abstracts Reference Service
Engineering Societies Library, US
American Society for Metals
Documents Librarian, The Center for Research Libraries, US

INFORMATION EXCHANGE AGREEMENT PARTNERS

Acquisitions Unit, Science Reference and Information Service, UK
Library - Exchange Desk, National Institute of Standards and Technology, US

SPARES (10 copies)

Total number of copies: 63

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION
DOCUMENT CONTROL DATA

1. PRIVACY MARKING/CAVEAT (OF DOCUMENT)

2. TITLE

Data Link Technology For A Portable Unmanned Aerial Vehicle

3. SECURITY CLASSIFICATION (FOR UNCLASSIFIED REPORTS THAT ARE LIMITED RELEASE USE (L) NEXT TO DOCUMENT CLASSIFICATION)

Document	(U)
Title	(U)
Abstract	(U)

4. AUTHOR(S)

Victor Kowalenko, Jane Phipps and Keith Cameron

5. CORPORATE AUTHOR

Aeronautical and Maritime Research Laboratory
PO Box 4331
Melbourne Vic 3001

6a. DSTO NUMBER	6b. AR NUMBER	6c. TYPE OF REPORT	7. DOCUMENT DATE
DSTO-RR-0087	AR-009-761	Research Report	November 1996

8. FILE NUMBER	9. TASK NUMBER	10. TASK SPONSOR	11. NO. OF PAGES	12. NO. OF REFERENCES
M1/9/156	76000B	DSTO	55	66

13. DOWNGRADING/DELIMITING INSTRUCTIONS

None

14. RELEASE AUTHORITY

Chief, Air Operations Division

15. SECONDARY RELEASE STATEMENT OF THIS DOCUMENT

Approved for public release

OVERSEAS ENQUIRIES OUTSIDE STATED LIMITATIONS SHOULD BE REFERRED THROUGH DOCUMENT EXCHANGE CENTRE, DIS NETWORK OFFICE, DEPT OF DEFENCE, CAMPBELL PARK OFFICES, CANBERRA ACT 2600

16. DELIBERATE ANNOUNCEMENT

No limitations

17. CASUAL ANNOUNCEMENT Yes

18. DEFTEST DESCRIPTORS

Data compression, data transmission, unmanned aerial vehicles, narrowband, high frequencies

19. ABSTRACT

This report examines data link requirements for a portable unmanned aerial vehicle. Crucial to the operation of such a data link is the development of suitable computer algorithms that are capable of significantly compressing and reconstructing image data in a timely manner for viewing at a remote station. As a consequence of the near real-time requirement, we investigate recent advances in lossy data compression techniques concentrating on transform coding techniques involving the discrete cosine transform, fractals and wavelets. At present the discrete cosine transform is available on a microprocessor chip and can offer acceptable reconstructed images close to real-time with compression ratios of up to 35:1, but other techniques promise even higher compression ratios and possibly a near real-time capability in the not too distant future.